GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 29, 2004, 17:37:45 ; Search time 13 Seconds
(without alignments)
1718.315 Million cell updates/sec

Title: US-10-024-652-2570
Perfect score: 2183
Sequence: 1 MAGSGAWKRLKSMLRKDDAP.....LQSYRQEVDRTCANCQSSSP 429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2179	99.8	429	1	ZNT4_HUMAN	O14863 homo sapien
2	2016.5	92.4	430	1	ZNT4_MOUSE	O35149 mus musculu
3	1990.5	91.2	430	1	ZNT4_RAT	O55174 rattus norv
4	747	34.2	359	1	ZNT2_RAT	Q62941 rattus norv
5	642	29.4	388	1	ZNT3_MOUSE	P97441 mus musculu
6	616	28.2	388	1	ZNT3_HUMAN	Q99726 homo sapien
7	386	17.7	313	1	ZITB_ECOLI	P75757 escherichia
8	385	17.6	311	1	ZITB_ECO57	Q8x400 escherichia
9	376	17.2	312	1	ZITB_SALTY	Q8zqt3 salmonella
10	372	17.0	312	1	ZITB_SALTI	Q8z8b6 salmonella
11	355	16.3	312	1	ZITB_YERPE	Q8zgy6 yersinia pe
12	353	16.2	316	1	CZCD_ALCEU	P13512 alcaligenes
13	353	16.2	316	1	CZCD_ALCSP	P94178 alcaligenes
14	343	15.7	439	1	COT1_YEAST	P32798 saccharomyc
15	342	15.7	442	1	ZRC1_YEAST	P20107 saccharomyc
16	289	13.2	507	1	ZNT1_HUMAN	Q9y6m5 homo sapien
17	284.5	13.0	507	1	ZNT1_RAT	Q62720 rattus norv
18	282.5	12.9	503	1	ZNT1_MOUSE	Q60738 mus musculu
19	230	10.5	732	1	YLH3_SCHPO	Q9hgq3 schizosacch
20	218	10.0	50	1	ZNT4_BOVIN	Q9ttf3 bos taurus
21	185	8.5	108	1	YGLB_BACST	P30540 bacillus st
22	171.5	7.9	283	1	Y449_METJA	Q57891 methanococc
23	160.5	7.4	290	1	YEAB_BACSU	P46348 bacillus su
24	148.5	6.8	310	1	YC63_SYNY3	P74068 synechocyst
25	137.5	6.3	251	1	YQR5_CAEEL	Q09308 caenorhabdi
26	136.5	6.3	510	1	YM43_YEAST	Q03218 saccharomyc
27	117.5	5.4	549	1	SMF2_YEAST	P38778 saccharomyc
28	107	4.9	567	1	YFJW_ECOLI	P52138 escherichia
29	107	4.9	1032	1	Y895_HAEIN	Q57124 haemophilus
30	105.5	4.8	1286	1	IRBP_BOVIN	P12661 bos taurus
31	103	4.7	301	1	P34_RICRI	P21559 rickettsia
32	100	4.6	309	1	RECS_MOUSE	Q8bjz3 mus musculu
33	100	4.6	547	1	YJDB_SALTY	P36555 salmonella
34	99.5	4.6	956	1	CLS3_MOUSE	Q99jh7 mus musculu
35	99	4.5	378	1	CKR7_MOUSE	P47774 mus musculu
36	99	4.5	532	1	SATT_MOUSE	O35874 mus musculu
37	99	4.5	956	1	CLS3_HUMAN	Q9bqt9 homo sapien
38	97.5	4.5	214	1	SP2M_BACSU	P37873 bacillus su
39	97.5	4.5	957	1	CLS3_RAT	Q8r553 rattus norv
40	97	4.4	705	1	MMLD_STRCO	Q9xa86 streptomyce
41	96.5	4.4	1205	1	PDS5_SCHPO	Q9hff5 schizosacch
42	96	4.4	300	1	P34_RICPR	Q9zcc5 rickettsia
43	96	4.4	366	1	PF2R_MOUSE	P43117 mus musculu
44	95	4.4	314	1	YFDV_ECOLI	P76519 escherichia
45	95	4.4	502	1	NU2C_MESVI	Q9muq6 mesostigma

ALIGNMENTS

RESULT 1
ZNT4_HUMAN
ID ZNT4_HUMAN STANDARD; PRT; 429 AA.
AC O14863;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc transporter 4 (ZnT-4).
GN SLC30A4 OR ZNT4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98016412; PubMed=9354792;
RA Huang L., Gitschier J.;
RT "A novel gene involved in zinc transport is deficient in the lethal
RT milk mouse.";
RL Nat. Genet. 17:292-297(1997).
CC -!- FUNCTION: Probably involved in zinc transport out of the
CC cytoplasm, may be by sequestration into an intracellular
CC compartment.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized in
CC endosomal vesicle membrane (Probable).
CC -!- DOMAIN: Contains a histidine-rich region which is a ligand for
CC zinc and an aspartate-rich region which is a potential ligand for
CC zinc (By similarity).
CC -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC ---

CC ---
DR EMBL; AF025409; AAB82561.1; -.
DR Genew; HGNC:11015; SLC30A4.
DR MIM; 602095; -.
DR GO; GO:0005385; F:zinc ion transporter activity; TAS.
DR InterPro; IPR002524; Cation_efflux.
DR Pfam; PF01545; Cation_efflux; 1.
DR TIGRFAMs; TIGR01297; CDF; 1.
KW Zinc transport; Transport; Transmembrane; Multigene family.
FT DOMAIN 1 113 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 114 134 POTENTIAL.
FT DOMAIN 135 143 VACUOLAR (POTENTIAL).
FT TRANSMEM 144 164 POTENTIAL.
FT DOMAIN 165 178 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 179 199 POTENTIAL.
FT DOMAIN 200 216 VACUOLAR (POTENTIAL).
FT TRANSMEM 217 237 POTENTIAL.

```
FT   DOMAIN      238    274       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    275    295       POTENTIAL.
FT   DOMAIN      296    310       VACUOLAR (POTENTIAL).
FT   TRANSMEM    311    331       POTENTIAL.
FT   DOMAIN      332    429       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       17     96       ASP-RICH (ACIDIC).
FT   DOMAIN      241    247       HIS-RICH.
SQ   SEQUENCE   429 AA;  47496 MW;  538992DFF8768739 CRC64;

  Query Match             99.8%;  Score 2179;  DB 1;  Length 429;
  Best Local Similarity   99.8%;  Pred. No. 1.5e-162;
  Matches  428;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFEFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 2
ZNT4_MOUSE
ID   ZNT4_MOUSE     STANDARD;      PRT;   430 AA.
AC   O35149; O35154;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Zinc transporter 4 (ZnT-4) (Lethal milk protein).
GN   SLC30A4 OR ZNT4 OR LM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6, and B6/CBAF1; TISSUE=Brain;
RX   MEDLINE=98016412; PubMed=9354792;
RA   Huang L., Gitschier J.;
RT   "A novel gene involved in zinc transport is deficient in the lethal
RT   milk mouse.";
RL   Nat. Genet. 17:292-297(1997).
CC   -!- FUNCTION: Probably involved in zinc transport out of the
CC       cytoplasm, may be by sequestration into an intracellular
CC       compartment.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized in
CC       endosomal vesicle membrane (Probable).
CC   -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in the
CC       brain and in mammary epithelial cell lines.
CC   -!- DOMAIN: Contains a histidine-rich region which is a ligand for
CC       zinc and an aspartate-rich region which is a potential ligand for
CC       zinc (By similarity).
CC   -!- DISEASE: Defects in SLC30A4 are the cause of the lethal milk (lm)
CC       phenotype. Mice with lm are defective in zinc transport into
CC       breast milk, due to a premature translation termination codon at
CC       position 297. Only homozygous mutant adults develop dermatitis,
CC       skin lesions, and hair loss due to a systemic zinc deficiency.
CC       However neonatal mice (of any genotype) suckled on homozygous
CC       mutant female also develop symptoms characteristic of nutritional
CC       zinc deficiency, including dermatitis, alopecia and stunted
CC       growth.
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF004099; AAB82413.1; -.
DR   EMBL; AF003747; AAB82593.1; -.
DR   EMBL; AF004098; AAB82412.1; -.
DR   EMBL; AF004097; AAB82411.1; -.
DR   EMBL; AF004100; AAB82414.1; -.
DR   MGD; MGI:1345282; Slc30a4.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Zinc transport; Transport; Transmembrane; Multigene family.
FT   DOMAIN        1    113       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    114    134       POTENTIAL.
FT   DOMAIN      135    143       VACUOLAR (POTENTIAL).
FT   TRANSMEM    144    164       POTENTIAL.
FT   DOMAIN      165    178       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    179    199       POTENTIAL.
FT   DOMAIN      200    216       VACUOLAR (POTENTIAL).
FT   TRANSMEM    217    237       POTENTIAL.
FT   DOMAIN      238    275       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    276    296       POTENTIAL.
FT   DOMAIN      297    311       VACUOLAR (POTENTIAL).
FT   TRANSMEM    312    332       POTENTIAL.
FT   DOMAIN      333    430       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       17     96       ASP-RICH (ACIDIC).
FT   DOMAIN      241    249       HIS-RICH.
FT   CONFLICT    426    426       Q -> H (IN REF. 1; AAB82414).
SQ   SEQUENCE   430 AA;  47791 MW;  E2311FA6265D15C6 CRC64;

  Query Match             92.4%;  Score 2016.5;  DB 1;  Length 430;
  Best Local Similarity   92.3%;  Pred. No. 7.1e-150;
  Matches  397;  Conservative   14;  Mismatches   16;  Indels    3;  Gaps    2;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||| |||||||:|||| ||||||||||||||  ||||||||||||||| ||||||
Db          1 MAGPGAWKRLKSLLRKDDTPLFLNDTSAFDFSDEVSDEGLSRFNKLRVVVADDDSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||| |||||||||||||||||:| ||||||:||:||||||| ||||||||
Db         61 PVNGAHPALQADDDSLLDQDLPLTNSQLSLKMDPCDNCSKRRELLKQRKVKTRLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              |||||||||||||:|||||||||||||||||||||||||||||||||||:||||||||||
Db        121 YLLFMIGELVGGYMANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTRRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||:|||:|||||||||||||||||||||||||:|||||||||||||||||||||
Db        181 VLSAMISVMLVYVLMGFLLYEAVQRTIHMNYEINGDVMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHS--HSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFK 298
```

```
                    | ||  ||||||||||:  |    ||||||||||||||||||||||||||||||||||||
Db        241 HHHSHAHSHSLPSNSPSMVSS-GHNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFK 299

Qy        299 PEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVE 358
              |||||||||||||:||||||||||||||||||||||||||||||||:|||||||||||
Db        300 PEYKIADPICTYIFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKESLMKIEDVYSVE 359

Qy        359 DLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVD 418
              ||||||||||||||||:|||||||||||||||| ||||||||||:||||||||||||
Db        360 DLNIWSLTSGKSTAIVHMQLIPGSSSKWEEVQSKAKHLLLNTFGMYKCTIQLQSYRQEVI 419

Qy        419 RTCANCQSSS 428
              ||||||||||
Db        420 RTCANCQSSS 429
```

```
RESULT 3
ZNT4_RAT
ID   ZNT4_RAT       STANDARD;      PRT;   430 AA.
AC   O55174;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Zinc transporter 4 (ZnT-4) (Dri 27 protein).
GN   SLC30A4 OR ZNT4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Brain;
RX   MEDLINE=20068535; PubMed=10600821;
RA   Murgia C., Vespignani I., Cerase J., Nobili F., Perozzi G.;
RT   "Cloning, expression, and vesicular localization of zinc transporter
RT   Dri 27/ZnT4 in intestinal tissue and cells.";
RL   Am. J. Physiol. 277:G1231-G1239(1999).
CC   -!- FUNCTION: Probably involved in zinc transport out of the
CC       cytoplasm, may be by sequestration into an intracellular
CC       compartment.
CC   -!- SUBUNIT: Mediates heterodimeric interactions with at least one
CC       specific partner.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized in
CC       endosomal vesicle membrane (Probable). In the polarized
CC       enterocytes, it is mainly localized in the basal cytoplasmic
CC       region.
CC   -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in brain
CC       and testes. Also expressed in small intestine, medulla, lung,
CC       kidney, stomach and colon. Expressed at lower level in other
CC       tissues.
CC   -!- DEVELOPMENTAL STAGE: Developmentaly regulated in the intestine.
CC   -!- INDUCTION: No change in response to zinc deprivation.
CC   -!- DOMAIN: Contains a histidine-rich region which is a ligand for
CC       zinc and an aspartate-rich region which is a potential ligand for
CC       zinc.
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y16774; CAA76372.1; -.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Zinc transport; Transport; Transmembrane; Multigene family.
FT   DOMAIN       1    113       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    114   134       POTENTIAL.
FT   DOMAIN      135   143       VACUOLAR (POTENTIAL).
FT   TRANSMEM    144   164       POTENTIAL.
FT   DOMAIN      165   178       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    179   199       POTENTIAL.
FT   DOMAIN      200   216       VACUOLAR (POTENTIAL).
FT   TRANSMEM    217   237       POTENTIAL.
FT   DOMAIN      238   275       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    276   296       POTENTIAL.
FT   DOMAIN      297   311       VACUOLAR (POTENTIAL).
FT   TRANSMEM    312   332       POTENTIAL.
FT   DOMAIN      333   430       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       17    96       ASP-RICH (ACIDIC).
FT   DOMAIN      241   249       HIS-RICH.
SQ   SEQUENCE   430 AA;  47702 MW;  F34CED3FA4FF05FB CRC64;

  Query Match             91.2%;  Score 1990.5;  DB 1;  Length 430;
  Best Local Similarity   90.7%;  Pred. No. 7.6e-148;
  Matches  390;  Conservative   19;  Mismatches   18;  Indels    3;  Gaps    2;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||| |||||||:|||||||||||||||||| ||  ||||||||||||||||| ||||||
Db          1 MAGPGAWKRLKSLLRKDDAPLFLNDTSAFDFLDEVSDEGLSRFNKLRVVVADDDSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||| |||||||||||:||||||||||:| |||||||:||||||| ||||||||
Db         61 PVNGAHPALQADDDSLLDQELPLTNSQLSLKMDPCDNCSKRRELLKQRKVKTRLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||:|||||||||||||||||||||||||||||||||:||||||||||
Db        121 YLLFMIGELVGGYMANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTRRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||:|||:|||||||||:||||||||||||:|||||||||||||||||||||||
Db        181 VLSAMISVMLVYVLMGFLLYEAMQRTIHMNYEINGDVMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHS--HSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFK 298
              | ||  ||||||||||:  |    :||||||||||||||||||||||||||||||||||
Db        241 HHHSHAHSHSLPSNSPSMVSS-GHSHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFK 299

Qy        299 PEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVE 358
              |||||||||||||:|||||||||||||||||||||||||||||||:|||||||||||
Db        300 PEYKIADPICTYIFSLLVAFTTLRIIWDTVVIILEGVPSHLNVDYIKESLMKIEDVYSVE 359

Qy        359 DLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVD 418
              ||||||||||:|||||:|||||||||||||||| ||||||||||:||:|||||||||
Db        360 DLNIWSLTSGKATAIVHMQLIPGSSSKWEEVQSKAKHLLLNTFGMYKCTVQLQSYRQEAT 419

Qy        419 RTCANCQSSS 428
              ||||||||||
Db        420 RTCANCQSSS 429
```

```
RESULT 4
ZNT2_RAT
ID   ZNT2_RAT       STANDARD;      PRT;   359 AA.
AC   Q62941;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Zinc transporter 2 (ZnT-2).
GN   SLC30A2 OR ZNT2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX   MEDLINE=96203098; PubMed=8617223;
RA   Palmiter R.D., Cole T.B., Findley S.D.;
RT   "ZnT-2, a mammalian protein that confers resistance to zinc by
RT   facilitating vesicular sequestration.";
RL   EMBO J. 15:1784-1791(1996).
```

```
CC   -!- FUNCTION: Involved in accumulation of zinc in endosomal/lysosomal
CC       vesicles.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized in the
CC       membrane of endosomal vesicles (Probable).
CC   -!- TISSUE SPECIFICITY: Intestine, kidney, seminal vesicles and
CC       testis.
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U50927; AAB02775.1; -.
DR   PIR; S70632; S70632.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Zinc transport; Transport; Transmembrane; Multigene family; Repeat.
FT   DOMAIN        1     56       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     57     77       POTENTIAL.
FT   DOMAIN       78     86       VACUOLAR (POTENTIAL).
FT   TRANSMEM     87    107       POTENTIAL.
FT   DOMAIN      108    123       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    124    144       POTENTIAL.
FT   DOMAIN      145    159       VACUOLAR (POTENTIAL).
FT   TRANSMEM    160    180       POTENTIAL.
FT   DOMAIN      181    207       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    208    228       POTENTIAL.
FT   DOMAIN      229    236       VACUOLAR (POTENTIAL).
FT   TRANSMEM    237    257       POTENTIAL.
FT   DOMAIN      258    291       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    292    312       POTENTIAL.
FT   DOMAIN      313    359       VACUOLAR (POTENTIAL).
FT   DOMAIN      183    192       5 X 2 AA APPROXIMATE REPEATS OF H-G.
SQ   SEQUENCE   359 AA;  39276 MW;  256A307021D1D197 CRC64;

  Query Match             34.2%;  Score 747;  DB 1;  Length 359;
  Best Local Similarity   42.0%;  Pred. No. 5.9e-51;
  Matches  148;  Conservative   78;  Mismatches  108;  Indels   18;  Gaps    4;

Qy         82 PLTNSQLSLKVDSCDNCSKQREI-----LKQRKVKARLTIAAVLYLLFMIGELVGGYIAN 136
              |:    : | | |   |  |::       :::: : :| :|: : |:|||||::|||:|
Db         20 PVNLPSVELAVQSNHYCHAQKDSGSHPNSEKQRARRKLYVASAICLVFMIGEIIGGYLAQ 79

Qy        137 SLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMG 196
              |||||||| |:||| ::::::| :||:||:  ||   ||: | |:| |::||| :::: |
Db         80 SLAIMTDAAHLLTDFASMLISLFSLWVSSRPATKTMNFGWQRAEILGALLSVLSIWVVTG 139

Qy        197 FLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTR 256
              |:| |||| |  :||| || ||||:   ||||:|||  |:|||| ||| ||   :|
Db        140 VLVYLAVQRLISGDYEIKGDTMLITSGCAVAVNIIMGLALHQSGHGHSHGHSHEDSS--- 196

Qy        257 GSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLV 316
                  | : :|||||:| :|||:||||||:||||| ||||||  |||||::||:||
Db        197 -------QQQQNPSVRAAFIHVVGDLLQSVGVLVAAYIIYFKPEYKYVDPICTFLFSILV 249

Qy        317 AFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHI 376
              ||  |: | :::::|| |  ::   :|  |: :: | ::  |:||:||  :    |||
Db        250 LGTTLTILRDVILVLMEGTPKGVDFTTVKNLLLSVDGVEALHSLHIWALTVAQPVLSVHI 309

Qy        377 QLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQSSS 428
              :     :  : |   |   |   | :  |||::|| ::: ::|  ||  |
Db        310 AIAQNVDA--QAVLKVARDRLQGKFNFHTMTIQIESYSEDM-KSCQECQGPS 358


RESULT 5
ZNT3_MOUSE
ID   ZNT3_MOUSE     STANDARD;      PRT;   388 AA.
AC   P97441; P97511;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Zinc transporter 3 (ZnT-3).
GN   SLC30A3 OR ZNT3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC   STRAIN=129; TISSUE=Brain;
RX   MEDLINE=97121493; PubMed=8962159;
RA   Palmiter R.D., Cole T.B., Quaife C.J., Findley S.D.;
RT   "ZnT-3, a putative transporter of zinc into synaptic vesicles.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:14934-14939(1996).
CC   -!- FUNCTION: Involved in accumulation of zinc in synaptic vesicles.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized in the
CC       membrane of synaptic vesicles (Probable).
CC   -!- TISSUE SPECIFICITY: Brain and testis. In the brain, most abundant
CC       in hippocampus and cerebral cortex. In the testis, expression is
CC       restricted to germ cells and is highest in pachytene spermatocytes
CC       and round spermatids.
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U76007; AAB39731.1; -.
DR   EMBL; U76009; AAB39733.1; ALT_SEQ.
DR   EMBL; U76008; AAB39733.1; JOINED.
DR   MGD; MGI:1345280; Slc30a3.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Zinc transport; Transport; Transmembrane; Multigene family.
FT   DOMAIN        1     75       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     76     96       POTENTIAL.
FT   DOMAIN       97    105       VACUOLAR (POTENTIAL).
FT   TRANSMEM    106    126       POTENTIAL.
FT   DOMAIN      127    145       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    146    166       POTENTIAL.
FT   DOMAIN      167    177       VACUOLAR (POTENTIAL).
FT   TRANSMEM    178    198       POTENTIAL.
FT   DOMAIN      199    235       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    236    256       POTENTIAL.
FT   DOMAIN      257    263       VACUOLAR (POTENTIAL).
FT   TRANSMEM    264    284       POTENTIAL.
FT   DOMAIN      285    388       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   388 AA;  41824 MW;  3CCDD0A37074EF41 CRC64;

  Query Match             29.4%;  Score 642;  DB 1;  Length 388;
  Best Local Similarity   35.0%;  Pred. No. 1e-42;
  Matches  144;  Conservative   85;  Mismatches  132;  Indels   50;  Gaps    9;

Qy         27 SAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPERPVNGAHPTLQADDDSLLDQDLPLTNS 86
              || | |   |  ||    :|: :  :     || |
Db         17 SARDRSSAGG--GL----RLKSLFTEPSEPLPEEP------------------------- 45

Qy         87 QLSLKVDSCDNCSKQ---REILKQRKVKAR--LTIAAVLYLLFMIGELVGGYIANSLAIM 141
                 |:  :  :| |    :  |   :|:||  |  |  :  :|| ||:||||:|:|||||
Db         46 --KLEGMAFHHCHKDPVPQSGLSPERVQARRQLYAACAVCFIFMAGEVVGGYLAHSLAIM 103

Qy        142 TDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYE 201
              ||| |:| |: ::: :| :||||::  |:  |||:|| | | |: ||: ::|: | |||
Db        104 TDAAHLLADIGSMLASLFSLWLSTRPATRTMTFGWHRSETLGALASVVSLWIVTGILLYL 163
```

```
Qy        202 AVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCE 261
              |  | :| :| |    ||:||:: |  |::| |:|:|:|  |||  :    :|      |
Db        164 AFLRLLHSDYHIEAGAMLLTASIAVCANLLMAFVLHQTGAPHSHGSTGAEYAP-----LE 218

Qy        262 RNHGQ----DSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVA 317
                ||       : :|||||| ||||:|| ||| |: :| |||:||:|||| |::||:
Db        219 EGHGYPMSLGNTSVRAAFVHVLGDLLQSFGVLAASILIYFKPQYKVADPISTFLFSICAL 278

Qy        318 FTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQ 377
               :|   : | :::::|| |  :  : ::: |: :  | :  ||::|:||    |  |:
Db        279 GSTAPTLRDVLLVLMEGAPRSVEFEPVRDTLLSVPGVRATHDLHLWALTLTYHVASAHLA 338

Qy        378 LIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQSSS 428
              :   |::  | | ::|:  | : ||   ||:|:: |: |: : |  ||  |
Db        339 I--DSTADPEAVLAEASSRLYSRFGFSSCTLQVEQYQPEMAQ-CLRCQEPS 386
```

RESULT 6
ZNT3_HUMAN
ID   ZNT3_HUMAN STANDARD; PRT; 388 AA.
AC   Q99726;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Zinc transporter 3 (ZnT-3).
GN   SLC30A3 OR ZNT3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Temporal cortex;
RX   MEDLINE=97121493; PubMed=8962159;
RA   Palmiter R.D., Cole T.B., Quaife C.J., Findley S.D.;
RT   "ZnT-3, a putative transporter of zinc into synaptic vesicles.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:14934-14939(1996).
CC   -!- FUNCTION: Involved in accumulation of zinc in synaptic vesicles
CC   (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized in the
CC   membrane of synaptic vesicles (Probable).
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC   transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; U76010; AAB39732.1; -.
DR   Genew; HGNC:11014; SLC30A3.
DR   MIM; 602878; -.
DR   GO; GO:0005768; C:endosome; TAS.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0005624; C:membrane fraction; TAS.
DR   GO; GO:0008021; C:synaptic vesicle; TAS.
DR   GO; GO:0015633; F:zinc porter activity; TAS.
DR   GO; GO:0006810; P:transport; TAS.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Zinc transport; Transport; Transmembrane; Multigene family.
FT   DOMAIN   1   75   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   76   96   POTENTIAL.
FT   DOMAIN   97   105   VACUOLAR (POTENTIAL).
FT   TRANSMEM   106   126   POTENTIAL.
FT   DOMAIN   127   145   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   146   166   POTENTIAL.
FT   DOMAIN   167   177   VACUOLAR (POTENTIAL).
FT   TRANSMEM   178   198   POTENTIAL.
FT   DOMAIN   199   235   CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   236   256   POTENTIAL.
FT   DOMAIN   257   264   VACUOLAR (POTENTIAL).
FT   TRANSMEM   265   285   POTENTIAL.
FT   DOMAIN   286   388   CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   388 AA; 41944 MW; 74B8E24F17CF4740 CRC64;

  Query Match   28.2%; Score 616; DB 1; Length 388;
  Best Local Similarity 37.2%; Pred. No. 1.1e-40;
  Matches 130; Conservative   81; Mismatches 132; Indels   6; Gaps   4;

```
Qy         79 QDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKAR--LTIAAVLYLLFMIGELVGGYIAN 136
              : ||  :  : :    |         |   :: ||  |  |  :  :|| ||:||||:|:
Db         39 EPLPEESKPVEMPFHHCHRDPLPPPGLTPERLHARRQLYAACAVCFVFMAGEVVGGYLAH 98

Qy        137 SLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMG 196
              |||||||| |:| |: ::: :| :||||::   |:  |||:|| | | |: ||: :::: |
Db         99 SLAIMTDAAHLLADVGSMMGSLFSLWLSTRPATRTMTFGWHRSETLGALASVVSLWMVTG 158

Qy        197 FLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTR 256
              ||| |  | :| :| | |  ||:||:: |  |::| |:|:|:|  |||          :|
Db        159 ILLYLAFVRLLHSDYHIEGGAMLLTASIAVCANLLMAFVLHQAGPPHSHGSRGAEYAPLE 218

Qy        257 GSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLV 316
              | |:     : :|||||| ||||:|| ||| |: :| |||:|| |||| |::||:
Db        219 -EGPEQPLPLGNTSVRAAFVHVLGDLLQSFGVLAASILIYFKPQYKAADPISTFLFSICA 277

Qy        317 AFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHI 376
              :|   : | : |::|| | ::  : ::: |: :  | :  :|::|:||     |  |:
Db        278 LGSTAPTLRDVLRILMEGTPRNVGFEPVRDTLLSVPGVRATHELHLWALTLTYHVASAHL 337

Qy        377 QLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQ 425
              :   |::  | | ::|:  | : ||   ||:|:: |: |: : |  ||
Db        338 AI--DSTADPEAVLAEASSRLYSRFGFSSCTLQVEQYQPEMAQ-CLRCQ 383
```

RESULT 7
ZITB_ECOLI
ID   ZITB_ECOLI STANDARD; PRT; 313 AA.
AC   P75757;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Zinc transporter zitB.
GN   ZITB OR B0752.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 12.7-28.0 min region on the linkage map.";
RL   DNA Res. 3:137-155(1996).

```
RN   [3]
RP   CHARACTERIZATION.
RC   STRAIN=K12 / W3110;
RX   MEDLINE=21336524; PubMed=11443104;
RA   Grass G., Fan B., Rosen B.P., Franke S., Nies D.H., Rensing C.;
RT   "ZitB (YbgR), a member of the cation diffusion facilitator family, is
RT   an additional zinc transporter in Escherichia coli.";
RL   J. Bacteriol. 183:4664-4667(2001).
CC   -!- FUNCTION: Involved in zinc efflux across the cytoplasmic membrane,
CC       thus reducing zinc accumulation in the cytoplasm and rendering
CC       bacteria more resistant to zinc. It may contribute to zinc
CC       homeostasis at low concentrations of zinc, whereas zntA is
CC       required for growth at more toxic concentrations.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- INDUCTION: By zinc.
CC   -!- MISCELLANEOUS: Appears to be selective for zinc, not conferring
CC       resistance to cobalt nor cadmium.
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000177; AAC73839.1; -.
DR   EMBL; D90714; BAA35414.1; -.
DR   PIR; H64810; H64810.
DR   EcoGene; EG13662; zitB.
DR   HAMAP; MF_00552; -; 1.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Transport; Zinc transport; Transmembrane; Complete proteome.
FT   TRANSMEM     21     41       POTENTIAL.
FT   TRANSMEM     48     68       POTENTIAL.
FT   TRANSMEM     90    110       POTENTIAL.
FT   TRANSMEM    122    142       POTENTIAL.
FT   TRANSMEM    160    180       POTENTIAL.
FT   TRANSMEM    182    202       POTENTIAL.
FT   DOMAIN      240    313       HIS-RICH.
SQ   SEQUENCE   313 AA;  34678 MW;  36740D67EABAC907 CRC64;

  Query Match            17.7%;  Score 386;  DB 1;  Length 313;
  Best Local Similarity  31.6%;  Pred. No. 7.1e-23;
  Matches   92;  Conservative   71;  Mismatches   94;  Indels   34;  Gaps    7;

Qy        124 FMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLS 183
              ||: |:|||::: |||:: || ||||| :|:: |||:  | : || | |||: ||  |:
Db         31 FMLVEVVGGFLSGSLALLADAGHMLTDTAALLFALLAVQFSRRPPTIRHTFGWLRLTTLA 90

Qy        184 AMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRH 243
              | :: : : ::   :::||::|       : | :|:  |  |:  |::  :||
Db         91 AFVNAIALVVITILIVWEAIER-FRTPRPVEGGMMMAIAVAGLLANILSFWLL------- 142

Qy        244 SHSHSLPSNSPTRGSGCERNHGQD--SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEY 301
                            :|| :  :| |||| :| ||||: ||| :||| || :   :
Db        143 -------------------HHGSEEKNLNVRAAALHVLGDLLGSVGAIIAALIIIW-TGW 182

Qy        302 KIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK-IEDVYSVEDL 360
                |||| : : |||| : :|:: |:|  :||| |  |::  :|  : : | :| :|  :
Db        183 TPADPILSILVSLLVLRSAWRLLKDSVNELLEGAPVSLDIAELKRRMCREIPEVRNVHHV 242

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              ::| :   |    :|:|:||      : :  :  | |:: : :   |||::
Db        243 HVW-MVGEKPVMTLHVQVIPPHDH--DALLDQIQHYLMDHYQIEHATIQME 290


RESULT 8
ZITB_ECO57
ID   ZITB_ECO57     STANDARD;      PRT;   311 AA.
AC   Q8X400; Q8X3F7;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Zinc transporter zitB.
GN   ZITB OR Z0922 OR ECS0780.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- FUNCTION: Involved in zinc efflux across the cytoplasmic membrane,
CC       thus reducing zinc accumulation in the cytoplasm and rendering
CC       bacteria more resistant to zinc. It may contribute to zinc
CC       homeostasis at low concentrations of zinc (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE005253; AAG55081.1; -.
DR   EMBL; AP002553; BAB34203.1; -.
DR   PIR; D90726; D90726.
DR   PIR; E85577; E85577.
DR   HAMAP; MF_00552; -; 1.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Transport; Zinc transport; Transmembrane; Complete proteome.
FT   TRANSMEM     19     39       POTENTIAL.
FT   TRANSMEM     46     66       POTENTIAL.
FT   TRANSMEM     88    108       POTENTIAL.
FT   TRANSMEM    120    140       POTENTIAL.
FT   TRANSMEM    158    178       POTENTIAL.
FT   TRANSMEM    180    200       POTENTIAL.
FT   DOMAIN      238    311       HIS-RICH.
FT   CONFLICT      1      2       MA -> MAHS (IN REF. 2).
SQ   SEQUENCE   311 AA;  34468 MW;  2D0F139E0C537587 CRC64;

  Query Match            17.6%;  Score 385;  DB 1;  Length 311;
  Best Local Similarity  31.3%;  Pred. No. 8.5e-23;
  Matches   91;  Conservative   72;  Mismatches   94;  Indels   34;  Gaps    7;
```

```
Qy        124 FMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLS 183
              ||: |::||::: |||:: || ||||| :|::  |||:  | : || | |||: ||  |:
Db         29 FMLVEVIGGFLSGSLALLADAGHMLTDTAALLFALLAVQFSRRPPTIRHTFGWLRLTTLA 88

Qy        184 AMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRH 243
              | :: : : ::   :::||::|       : | :|:  |  |:  |::  :||
Db         89 AFVNAIALVVITILIVWEAIER-FRTPRPVEGGMMMAIAVAGLLANILSFWLL------- 140

Qy        244 SHSHSLPSNSPTRGSGCERNHGQD--SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEY 301
                              :|| :  :| |||| :| ||||: ||| :||| || :   :
Db        141 ------------------HHGSEEKNLNVRAAALHVLGDLLGSVGAIIAALIIIW-TGW 180

Qy        302 KIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK-IEDVYSVEDL 360
                |||| : : |||| : :|:: |:|  :||| |  |::  :|  : : | :| :|  :
Db        181 TPADPILSILVSLLVLRSAWRLLKDSVNELLEGAPVSLDIAELKRRMCREIPEVRNVHHV 240

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              ::| :   |    :|:|:||      : :  :  | |:: : :   ||::
Db        241 HVW-MVGEKPVMTLHVQVIPPHDH--DALLDQIQHYLMDHYQIEHATIQME 288


RESULT 9
ZITB_SALTY
ID   ZITB_SALTY     STANDARD;      PRT;   312 AA.
AC   Q8ZQT3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Zinc transporter zitB.
GN   ZITB OR STM0758.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX   MEDLINE=21534948; PubMed=11677609;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
CC   -!- FUNCTION: Involved in zinc efflux across the cytoplasmic membrane,
CC       thus reducing zinc accumulation in the cytoplasm and rendering
CC       bacteria more resistant to zinc. It may contribute to zinc
CC       homeostasis at low concentrations of zinc (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008731; AAL19697.1; -.
DR   StyGene; SG?????; zitB.
DR   HAMAP; MF_00552; -; 1.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Transport; Zinc transport; Transmembrane; Complete proteome.
FT   TRANSMEM     21     41       POTENTIAL.
FT   TRANSMEM     48     68       POTENTIAL.
FT   TRANSMEM     90    110       POTENTIAL.
FT   TRANSMEM    123    143       POTENTIAL.
FT   TRANSMEM    164    184       POTENTIAL.
FT   DOMAIN      240    312       HIS-RICH.
SQ   SEQUENCE   312 AA;  34461 MW;  2521FF85B2B7B3A6 CRC64;

  Query Match             17.2%;  Score 376;  DB 1;  Length 312;
  Best Local Similarity   31.0%;  Pred. No. 4.3e-22;
  Matches   93;  Conservative   77;  Mismatches  100;  Indels   30;  Gaps    8;

Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRF 172
              ||  | ::   ||: |:||| :: |||:: || ||||| :|::  |||:  | : || |
Db         20 RLLFAFIVTAGFMLLEVVGGILSGSLALLADAGHMLTDAAALLFALLAVQFSRRPPTVRH 79

Qy        173 TFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIM 232
              |||: ||  |:| :: : : ::   :::||::|  :    : |::|:: |  |:  |:
Db         80 TFGWLRLTTLAAFVNAIALVVITLLIVWEAIER-FYTPRPVAGNLMMVIAVAGLLANLFA 138

Qy        233 GFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAA 292
              ::|:                  |||  |:|     | |||| :| :|||: ||| ::||
Db        139 FWILH-----------------RGSD-EKN-----LNVRAAALHVMGDLLGSVGAIVAA 174

Qy        293 YIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK-I 351
              || :   :  |||| : : |:||  : :|:: |:|  :||| |  |::: ::  | : |
Db        175 LIIIW-TGWTPADPILSILVSVLVLRSAWRLLKDSVNELLEGAPVSLDINALQRHLSREI 233

Qy        352 EDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              :| :|  :::| :   |    :| |:||      : :  :  |:: : :   |||::
Db        234 PEVRNVHHVHVW-MVGEKPVMTLHAQVIPPHDH--DALLERIQDFLMHEYHIAHATIQME 290


RESULT 10
ZITB_SALTI
ID   ZITB_SALTI     STANDARD;      PRT;   312 AA.
AC   Q8Z8B6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Zinc transporter zitB.
GN   ZITB OR STY0799 OR T2120.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Ty2 / ATCC 700931;
RX   MEDLINE=22531367; PubMed=12644504;
RA   Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA   Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT   "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT   and CT18.";
RL   J. Bacteriol. 185:2330-2337(2003).
CC   -!- FUNCTION: Involved in zinc efflux across the cytoplasmic membrane,
CC       thus reducing zinc accumulation in the cytoplasm and rendering
CC       bacteria more resistant to zinc. It may contribute to zinc
CC       homeostasis at low concentrations of zinc (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL627268; CAD05215.1; -.
DR   EMBL; AE016841; AAO69737.1; -.
DR   HAMAP; MF_00552; -; 1.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Transport; Zinc transport; Transmembrane; Complete proteome.
FT   TRANSMEM     21     41       POTENTIAL.
FT   TRANSMEM     48     68       POTENTIAL.
FT   TRANSMEM     90    110       POTENTIAL.
FT   TRANSMEM    123    143       POTENTIAL.
FT   TRANSMEM    164    184       POTENTIAL.
FT   DOMAIN      240    312       HIS-RICH.
SQ   SEQUENCE   312 AA;  34531 MW;  EA29EF94B3A18736 CRC64;

  Query Match             17.0%;  Score 372;  DB 1;  Length 312;
  Best Local Similarity   30.7%;  Pred. No. 8.8e-22;
  Matches   92;  Conservative   77;  Mismatches  101;  Indels   30;  Gaps    8;

Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRF 172
              ||   | ::   ||: |:||| :: |||:: || ||||| :|::  || :  | : || |
Db         20 RLLFAFIVTAGFMLLEVVGGILSGSLALLADAGHMLTDAAALLFALLVVQFSRRPPTVRH 79

Qy        173 TFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIM 232
              |||: ||  |:| :: : : ::   :::||::|  :    : |::|:: |  |:  |:
Db         80 TFGWLRLTTLAAFVNAIALVVITLLIVWEAIER-FYTPRPVAGNLMMVIAVAGLLANLFA 138

Qy        233 GFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAA 292
               ::|:                 |||  |:|     | |||| :| :|||: ||| ::||
Db        139 FWILH-----------------RGSD-EKN-----LNVRAAALHVMGDLLGSVGAIVAA 174

Qy        293 YIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK-I 351
              || :   :  |||| : : |:||  : :|:: |:|  :||| |  |::: ::  | : |
Db        175 LIIIW-TGWTPADPILSILVSVLVLRSAWRLLKDSVNELLEGAPVSLDINALQRHLSREI 233

Qy        352 EDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              :| :|  :::| :   |    :| |:||      : :  :    |:: : :   |||::
Db        234 PEVRNVHHVHVW-MVGEKPVMTLHAQVIPPHDH--DALLERIQDFLMHEYHIAHATIQME 290


RESULT 11
ZITB_YERPE
ID   ZITB_YERPE     STANDARD;      PRT;   312 AA.
AC   Q8ZGY6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Zinc transporter zitB.
GN   ZITB OR YPO1129 OR Y3050.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / Biovar Orientalis;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KIM5 / Biovar Mediaevalis;
RX   MEDLINE=22137863; PubMed=12142430;
RA   Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA   Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA   Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA   Perry R.D.;
RT   "Genome sequence of Yersinia pestis KIM.";
RL   J. Bacteriol. 184:4601-4611(2002).
CC   -!- FUNCTION: Involved in zinc efflux across the cytoplasmic membrane,
CC       thus reducing zinc accumulation in the cytoplasm and rendering
CC       bacteria more resistant to zinc. It may contribute to zinc
CC       homeostasis at low concentrations of zinc (By similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ414146; CAC89972.1; -.
DR   EMBL; AE013906; AAM86600.1; -.
DR   PIR; AI0138; AI0138.
DR   HAMAP; MF_00552; -; 1.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Transport; Zinc transport; Transmembrane; Complete proteome.
FT   TRANSMEM     16     36       POTENTIAL.
FT   TRANSMEM     40     60       POTENTIAL.
FT   TRANSMEM     81    101       POTENTIAL.
FT   TRANSMEM    117    137       POTENTIAL.
FT   TRANSMEM    153    173       POTENTIAL.
FT   TRANSMEM    177    197       POTENTIAL.
FT   DOMAIN      235    312       HIS-RICH.
SQ   SEQUENCE   312 AA;  34827 MW;  B727E66194A66705 CRC64;

  Query Match             16.3%;  Score 355;  DB 1;  Length 312;
  Best Local Similarity   27.5%;  Pred. No. 1.9e-20;
  Matches   85;  Conservative   77;  Mismatches  117;  Indels   30;  Gaps    6;

Qy        104 ILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWL 163
              |  |      || ||  :  |||: | :||::: |||:: || ||||| :|: : |:|:
Db          6 IFSQDSNSKRLLIAFAITTLFMVTEAIGGWLSGSLALLADAGHMLTDSAALFIALMAVHF 65

Qy        164 SSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAA 223
              | : |  | ||: ||  |:| ::   : :::  :::||| |     :|: |  ||| |
Db         66 SQRKPDPRHTFGYLRLTTLAAFVNAAALLLIVILIVWEAVHR-FFSPHEVMGTPMLIIAI 124

Qy        224 VGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLV 283
              |:  |:   ::|:                        :   : :: |||| :| | ||:
Db        125 AGLLANIFCFWILH----------------------KGEEEKNINVRAAALHVLSDLL 160

Qy        284 QSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDY 343
              ||| :||| :|    :   ||| : : |:|:  : :|:: ::   :||| |  ::::
Db        161 GSVGAMIAAIVI-LTTGWTPIDPILSVLVSVLILRSAWRLLKESFHELLEGAPQEIDINK 219

Qy        344 I-KEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFG 402
              : |:    | :| ::  :::| :   :   : | |:||      : :  :    ||: :
Db        220 LRKDLCTNIYEVRNIHHVHLWQVGEQRLMTL-HAQVIPPLDH--DALLQRIQDYLLHHYR 276

Qy        403 MYRCTIQLQ 411
              :   |:|::
Db        277 ISHATVQME 285
```

```
RESULT 12
CZCD_ALCEU
ID   CZCD_ALCEU     STANDARD;      PRT;   316 AA.
AC   P13512;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Cobalt-zinc-cadmium resistance protein czcD (Cation efflux system
DE   protein czcD).
GN   CZCD.
OS   Alcaligenes eutrophus (Ralstonia eutropha).
OG   Plasmid pMOL30.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=510;
RN   [1]
RP   PRELIMINARY SEQUENCE FROM N.A.
RC   STRAIN=CH34;
RX   MEDLINE=90017477; PubMed=2678100;
RA   Nies D.H., Nies A., Chu L., Silver S.;
RT   "Expression and nucleotide sequence of a plasmid-determined divalent
RT   cation efflux system from Alcaligenes eutrophus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:7351-7355(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CH34;
RX   MEDLINE=97197194; PubMed=9044283;
RA   van der Lelie D., Schwuchow T., Schwidetzky T., Wuertz S.,
RA   Baeyens W., Mergeay M., Nies D.H.;
RT   "Two-component regulatory system involved in transcriptional control
RT   of heavy-metal homoeostasis in Alcaligenes eutrophus.";
RL   Mol. Microbiol. 23:493-503(1997).
CC   -!- FUNCTION: Necessary for activation of the czc determinant.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- INDUCTION: By zinc.
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X98451; CAA67085.1; -.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Transport; Zinc transport; Cadmium resistance; Cadmium; Cobalt;
KW   Transmembrane; Plasmid.
FT   TRANSMEM     17     37       POTENTIAL.
FT   TRANSMEM     47     67       POTENTIAL.
FT   TRANSMEM     82    102       POTENTIAL.
FT   TRANSMEM    115    135       POTENTIAL.
FT   TRANSMEM    152    172       POTENTIAL.
FT   TRANSMEM    174    194       POTENTIAL.
SQ   SEQUENCE   316 AA;  33707 MW;  2977C1B4AEE9600F CRC64;

  Query Match            16.2%;  Score 353;  DB 1;  Length 316;
  Best Local Similarity  30.8%;  Pred. No. 2.7e-20;
  Matches   92;  Conservative   67;  Mismatches  110;  Indels   30;  Gaps    6;

Qy        114 LTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFT 173
              | ||  |   |:| |:||| :  |||:::|| |||||  |: : | |: :: :   |: |
Db         17 LKIALALTGTFLIAEVVGGVMTKSLALISDAAHMLTDTVALAIALAAIAIAKRPADKKRT  76

Qy        174 FGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMG 233
              ||::| |:|:|  : ||:: :  ::|||  | :    :|    | : | :|: :|:|
Db         77 FGYYRFEILAAAFNALLLFGVAIYILYEAYLR-LKSPPQIESTGMFVVAVLGLIINLISM 135

Qy        234 FLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAY 293
              :|: ||                          || |: |::    ||: ||||:  |
Db        136 RMLS-SGQ----------------------SSSLNVKGAYLEVWSDLLGSVGVIAGAI 170

Qy        294 IIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIED 353
              |||| :  |    :  | |  |: ::  :: ::|||||  :::  ::: ::
Db        171 IIRF-TGWAWVDSAIAVLIGLWVLPRTWILLKSSLNVLLEGVPDDVDLAEVEKQILATPG 229

Qy        354 VYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWE-EVQSKANHLLLNTFGMYRCTIQLQ 411
              | |  ||:||:|||||::  ||:  :  ::   | ||  :   :| : | :   ||| :
Db        230 VKSFHDLHIWALTSGKASLTVHV--VNDTAVNPEMEVLPELKQMLADKFDITHVTIQFE 286


RESULT 13
CZCD_ALCSP
ID   CZCD_ALCSP     STANDARD;      PRT;   316 AA.
AC   P94178;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cation efflux system protein czcD.
GN   CZCD.
OS   Alcaligenes sp. (strain CT14).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Alcaligenes.
OX   NCBI_TaxID=512;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96219090; PubMed=8829543;
RA   Kunito T., Kusano T., Oyaizu H., Senoo K., Kanazawa S.,
RA   Matsumoto S.;
RT   "Cloning and sequence analysis of czc genes in Alcaligenes sp. strain
RT   CT14.";
RL   Biosci. Biotechnol. Biochem. 60:699-704(1996).
CC   -!- FUNCTION: Necessary for activation of the czc determinant.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D67044; BAA11062.1; -.
DR   PIR; JC4701; JC4701.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Transport; Transmembrane; Zinc transport; Cadmium; Cobalt.
FT   TRANSMEM     17     37       POTENTIAL.
FT   TRANSMEM     47     67       POTENTIAL.
FT   TRANSMEM     82    102       POTENTIAL.
FT   TRANSMEM    115    135       POTENTIAL.
FT   TRANSMEM    152    172       POTENTIAL.
FT   TRANSMEM    174    194       POTENTIAL.
SQ   SEQUENCE   316 AA;  33741 MW;  D677C1B4A1E96462 CRC64;

  Query Match            16.2%;  Score 353;  DB 1;  Length 316;
  Best Local Similarity  30.8%;  Pred. No. 2.7e-20;
  Matches   92;  Conservative   67;  Mismatches  110;  Indels   30;  Gaps    6;

Qy        114 LTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFT 173
              | ||  |   |:| |:||| :  |||:::|| |||||  |: : | |: :: :   |: |
Db         17 LKIALALTGTFLIAEVVGGVMTKSLALISDAAHMLTDTVALAIALAAIAIAKRPADKKRT  76

Qy        174 FGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMG 233
              ||::| |:|:|  : ||:: :  ::|||  | :    :|    | : | :|: :|:|
Db         77 FGYYRFEILAAAFNALLLFGVAIYILYEAYLR-LKSPPQIESTGMFVVAVLGLIINLISM 135
```

```
Qy        234 FLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAY 293
              :|: ||                                 || |: |::    ||: ||||:  |
Db        136 RMLS-SGQ-----------------------SSSLNVKGAYLEVWSDLLGSVGVIAGAI 170

Qy        294 IIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIED 353
              ||||   :   |     :  | |   |: ::  :: ::|||||  :::  ::: ::
Db        171 IIRF-TGWAWVDSAIAVLIGLWVLPRTWFLLKSSLNVLLEGVPDDVDLAEVEKQILATPG 229

Qy        354 VYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWE-EVQSKANHLLLNTFGMYRCTIQLQ 411
              | |  ||:||:|||||::  ||:  :  ::   | ||  :   :| : | :   ||| :
Db        230 VKSFHDLHIWALTSGKASLTVHV--VNDTAVNPEMEVLPELKQMLADKFDITHVTIQFE 286


RESULT 14
COT1_YEAST
ID   COT1_YEAST     STANDARD;      PRT;   439 AA.
AC   P32798;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Cobalt uptake protein COT1.
GN   COT1 OR YOR316C OR O6131.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92375034; PubMed=1508175;
RA   Conklin D.S., McMaster J.A., Culbertson M.R., Kung C.;
RT   "COT1, a gene involved in cobalt accumulation in Saccharomyces
RT   cerevisiae.";
RL   Mol. Cell. Biol. 12:3678-3688(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / FY1679;
RX   MEDLINE=97051589; PubMed=8896266;
RA   Pearson B.M., Hernando Y., Payne J., Wolf S.S., Kalogeropoulos A.,
RA   Schweizer M.;
RT   "Sequencing of a 35.71 kb DNA segment on the right arm of yeast
RT   chromosome XV reveals regions of similarity to chromosomes I and
RT   XIII.";
RL   Yeast 12:1021-1031(1996).
CC   -!- FUNCTION: PROBABLY RESPONSIBLE FOR THE UPTAKE OF COBALT IONS. IT
CC       APPEARS TO ACT IN A DOSAGE-DEPENDENT MANNER TO COUNTERACT THE
CC       ADVERSE EFFECTS OF COBALT IONS ON CELLS. IT MAY PARTICIPATE IN
CC       THE REGULATION OF COBALT LEVELS UNDER NORMAL PHYSIOLOGICAL
CC       CONDITIONS AND MAY BE IMPORTANT IN THE SUPPLY OF METAL THAT IS
CC       REQUIRED FOR METALLOENZYME OR COFACTOR SYNTHESIS. IT REDUCES THE
CC       TOXICITY OF COBALT AND RHODIUM IONS. OTHER COMPONENTS RESPONSIBLE
CC       FOR COBALT TRANSPORT EXIST.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MEMBRANE. ANOTHER POSSIBILITY
CC       EXISTS THAT IT IS ASSOCIATED WITH ANOTHER UNIDENTIFIED MEMBRANE
CC       THAT HAS BEEN ENRICHED IN THE MITOCHONDRIAL MEMBRANE FRACTIONS.
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M88252; AAA74884.1; -.
DR   EMBL; X90565; CAA62171.1; -.
DR   EMBL; Z75224; CAA99636.1; -.
DR   PIR; S58327; S58327.
DR   GermOnline; 143904; -.
DR   SGD; S0005843; COT1.
DR   GO; GO:0000324; C:vacuole (sensu Fungi); IDA.
DR   GO; GO:0015087; F:cobalt ion transporter activity; IMP.
DR   GO; GO:0005385; F:zinc ion transporter activity; IGI.
DR   GO; GO:0006824; P:cobalt ion transport; IMP.
DR   GO; GO:0006882; P:zinc ion homeostasis; IMP.
DR   GO; GO:0006829; P:zinc ion transport; IGI.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Transport; Cobalt; Mitochondrion; Transmembrane.
FT   TRANSMEM     10     27       POTENTIAL.
FT   TRANSMEM     43     60       POTENTIAL.
FT   TRANSMEM     78    100       POTENTIAL.
FT   TRANSMEM    114    133       POTENTIAL.
FT   TRANSMEM    244    265       POTENTIAL.
FT   TRANSMEM    279    295       POTENTIAL.
FT   DOMAIN      140    148       HIS-RICH; COULD BE INVOLVED IN
FT                                COORDINATION OF COBALT IONS.
FT   DOMAIN      163    169       HIS-RICH; COULD BE INVOLVED IN
FT                                COORDINATION OF COBALT IONS.
FT   CONFLICT    227    227       G -> E (IN REF. 1).
FT   CONFLICT    333    334       HI -> RV (IN REF. 1).
FT   CONFLICT    424    424       I -> V (IN REF. 1).
SQ   SEQUENCE   439 AA;  48154 MW;  AC88AAA5F2EE4AED CRC64;

  Query Match             15.7%;  Score 343;  DB 1;  Length 439;
  Best Local Similarity   25.3%;  Pred. No. 2.4e-19;
  Matches  101;  Conservative   77;  Mismatches  123;  Indels   98;  Gaps   12;

Qy        106 KQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALW--- 162
              || |: : | :  | : :    |:  ||:::|||:: |: ||| |   ||  ::|||
Db          6 KQVKIISLLLLDTVFFGI----EITTGYLSHSLALIADSFHMLND---IISLVVALWAVN 58

Qy        163 -LSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLIT 221
               :::|   :|:|: | |:| |:|: : :  |   :| ||:|| |       |   :|
Db         59 VAKNRNPDSTYTYGWKRAEILGALINAVFLIALCVSILIEALQRIIAPPVIENPKFVLYV 118

Qy        222 AAVGVAVNVIMGFLLN----QSGHRHSHSHS---------LPS----------------- 251
              |:  | :  || :    : || | |||           :||
Db        119 GVAGLISNTVGLFLFHDNDQEHGHGHGHSHGGIFADHEMHMPSSHTHTHAHVDGIENTTP 178

Qy        252 --------------------NSPTR----------------------GSGCERNHGQDS 268
                                 |  ||                     |:  | |  : |
Db        179 MDSTDNISEIMPNAIVDSFMNENTRLLTPENASKTPSYSTSSHTIASGGNYTEHNKRKRS 238

Qy        269 LAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYK---IADPICTYVFSLLVAFTTFRIIW 325
              | :   |:| ||| : ::||:::|: | :| :|       ||: : : : ::  :   :
Db        239 LNMHGVFLHVLGDALGNIGVMLSAFFI-WKTDYSWKYYTDPLVSLIITGIIFSSALPLSC 297

Qy        326 DTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSK 385
              |:|:  || |: | ::  |:||  : :: | :||:||      | :||||   |  :
Db        298 KASKILLQATPSTLSGDQVEGDLLKIPGIIAIHDFHIWNLTESIFIASLHIQL-DISPEQ 356

Qy        386 WEE----VQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              : :    |:||     |: :|::  |:| :   :|| |
Db        357 FTDLAKIVRSK-----LHRYGIHSATLQPEFITREVTST 390


RESULT 15
ZRC1_YEAST
ID   ZRC1_YEAST     STANDARD;      PRT;   442 AA.
AC   P20107;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Zinc/cadmium resistance protein.
GN   ZRC1 OR YMR243C OR YM9408.05C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DKD-5D-H;
```

```
RX   MEDLINE=90136503; PubMed=2693940;
RA   Kamizono A., Nishizawa M., Teranishi Y., Murata K., Kimura A.;
RT   "Identification of a gene conferring resistance to zinc and cadmium
RT   ions in the yeast Saccharomyces cerevisiae.";
RL   Mol. Gen. Genet. 219:161-167(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313268; PubMed=9169872;
RA   Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA   Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA   Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA   Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT   XIII.";
RL   Nature 387:90-93(1997).
CC   -!- FUNCTION: Probably responsible for the uptake of zinc and cadmium
CC       ions.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC       (Potential).
CC   -!- SIMILARITY: Belongs to the cation diffusion facilitator family of
CC       transporters (CDF, TC 2.A.4). SLC30A subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17537; CAB56542.1; -.
DR   EMBL; Z48756; CAA88653.1; -.
DR   PIR; S56057; S56057.
DR   GermOnline; 142918; -.
DR   SGD; S0004856; ZRC1.
DR   GO; GO:0000324; C:vacuole (sensu Fungi); IDA.
DR   GO; GO:0005385; F:zinc ion transporter activity; IMP.
DR   GO; GO:0006824; P:cobalt ion transport; NAS.
DR   GO; GO:0006749; P:glutathione metabolism; NAS.
DR   GO; GO:0006882; P:zinc ion homeostasis; IMP.
DR   GO; GO:0006829; P:zinc ion transport; NAS.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Transport; Zinc transport; Cadmium resistance; Mitochondrion; Cadmium;
KW   Transmembrane; Glycoprotein.
FT   TRANSMEM     10     26       POTENTIAL.
FT   TRANSMEM     42     58       POTENTIAL.
FT   TRANSMEM     80     96       POTENTIAL.
FT   TRANSMEM    113    129       POTENTIAL.
FT   TRANSMEM    240    256       POTENTIAL.
FT   TRANSMEM    270    286       POTENTIAL.
FT   DOMAIN      141    145       HIS-RICH; COULD BE INVOLVED IN
FT                                COORDINATION OF ZINC OR CADMIUM IONS.
FT   DOMAIN      163    167       HIS-RICH; COULD BE INVOLVED IN
FT                                COORDINATION OF ZINC OR CADMIUM IONS.
FT   DOMAIN      216    220       HIS-RICH; COULD BE INVOLVED IN
FT                                COORDINATION OF ZINC OR CADMIUM IONS.
FT   CONFLICT    414    414       R -> S (IN REF. 1).
FT   CONFLICT    417    417       T -> I (IN REF. 1).
SQ   SEQUENCE   442 AA;  48344 MW;  CBE6AA9716C0CEBE CRC64;

  Query Match             15.7%;  Score 342;  DB 1;  Length 442;
  Best Local Similarity   26.1%;  Pred. No. 2.9e-19;
  Matches   98;  Conservative   73;  Mismatches  123;  Indels   82;  Gaps   12;

Qy        106 KQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALW--- 162
              |: :: : ||:  | :||    |:  ||:::|||:: |: ||| |   ||  |:|||
Db          5 KELRIISLLTLDTVFFLL----EITIGYMSHSLALIADSFHMLND---IISLLVALWAVD 57

Qy        163 -LSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLIT 221
               :: |  ::|:|: | |:| |:|: : :  |   :: ||:|| |      |  ::|
```

```
Db         58 VAKNRGPDAKYTYGWKRAEILGALINAVFLIALCFSIMIEALQRLIEPQEIQNPRLVLYV 117

Qy        222 AAVGVAVNVIMGFLLNQSG----HRHS-----------------HSHS---------- 248
                |:  ||:  || :  |    | ||                 ||||
Db        118 GVAGLISNVVGLFLFHDHGSDSLHSHSHGSVESGNNDLDIESNATHSHSHASLPNDNLAI 177

Qy        249 ----------------LPSNSPTRGSGCER---NH-----------GQDSLAVRAAFV 276
                             || :  | |  :  ||            |  || :   |:
Db        178 DEDAISSPGPSGQIGEVLPQSVVNRLSNESQPLLNHDDHDHSHESKKPGHRSLNMHGVFL 237

Qy        277 HALGDLVQSVGVLIAAYIIRFKPEYK---IADPICTYVFSLLVAFTTFRIIWDTVVIILE 333
              | ||| : ::||: ||  | :| ||     :||| : : ::::  :   :      |:|:
Db        238 HVLGDALGNIGVIAAALFI-WKTEYSWRYYSDPIVSLIITIIIFSSALPLSRRASRILLQ 296

Qy        334 GVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKA 393
               || :: | |:  :: :  | :| | ::|:||     | :|:| |  :  |:   :|
Db        297 ATPSTISADQIQREILAVPGVIAVHDFHVWNLTESIYIASIHVQ-IDCAPDKFMS-SAKL 354

Qy        394 NHLLLNTFGMYRCTIQ 409
                : :  |::  |:|
Db        355 IRKIFHQHGIHSATVQ 370
```

Search completed: July 29, 2004, 17:43:32
Job time : 14 secs

Blank sheet

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 29, 2004, 17:43:15 ; Search time 46 Seconds
(without alignments)
2925.434 Million cell updates/sec

Title: US-10-024-652-2570
Perfect score: 2183
Sequence: 1 MAGSGAWKRLKSMLRKDDAP.....LQSYRQEVDRTCANCQSSSP 429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2183	100.0	429	15	US-10-024-652-2570	Sequence 2570, Ap
2	2183	100.0	429	15	US-10-024-652-2572	Sequence 2572, Ap
3	2183	100.0	429	15	US-10-024-652-2575	Sequence 2575, Ap
4	2183	100.0	429	15	US-10-024-652-2576	Sequence 2576, Ap
5	2183	100.0	429	15	US-10-024-652-2577	Sequence 2577, Ap
6	2183	100.0	429	15	US-10-024-652-2578	Sequence 2578, Ap
7	2183	100.0	429	15	US-10-295-027-544	Sequence 544, App
8	2183	100.0	438	10	US-09-957-708-27	Sequence 27, Appl
9	2179	99.8	429	14	US-10-345-680-53	Sequence 53, Appl
10	2179	99.8	429	15	US-10-024-652-2574	Sequence 2574, Ap
11	2179	99.8	429	15	US-10-024-652-2579	Sequence 2579, Ap
12	2176	99.7	428	15	US-10-024-652-2580	Sequence 2580, Ap
13	2166.5	99.2	430	15	US-10-012-697-1545	Sequence 1545, Ap
14	1990.5	91.2	429	15	US-10-024-652-2581	Sequence 2581, Ap
15	747	34.2	359	10	US-09-954-342-37	Sequence 37, Appl
16	723	33.1	356	12	US-10-114-270-4	Sequence 4, Appli
17	722.5	33.1	372	16	US-10-380-727-4	Sequence 4, Appli
18	718.5	32.9	372	14	US-10-162-012-40	Sequence 40, Appl
19	718.5	32.9	372	15	US-10-162-102-40	Sequence 40, Appl
20	679	31.1	323	12	US-10-114-270-6	Sequence 6, Appli
21	653.5	29.9	369	10	US-09-954-342-8	Sequence 8, Appli
22	648.5	29.7	369	10	US-09-954-342-12	Sequence 12, Appl
23	648	29.7	320	9	US-09-872-153-22	Sequence 22, Appl
24	648	29.7	320	14	US-10-176-306-20	Sequence 20, Appl
25	642	29.4	388	10	US-09-954-342-38	Sequence 38, Appl
26	617.5	28.3	422	10	US-09-954-342-10	Sequence 10, Appl
27	616	28.2	388	10	US-09-954-342-39	Sequence 39, Appl
28	616	28.2	388	12	US-10-369-022-44	Sequence 44, Appl
29	616	28.2	388	14	US-10-176-306-52	Sequence 52, Appl
30	557.5	25.5	234	12	US-09-826-734-200	Sequence 200, App
31	551	25.2	415	12	US-10-425-114-60835	Sequence 60835, A
32	539	24.7	448	12	US-10-425-114-61150	Sequence 61150, A
33	536.5	24.6	418	16	US-10-437-963-111869	Sequence 111869,
34	530	24.3	274	15	US-10-104-047-3913	Sequence 3913, Ap
35	505.5	23.2	399	12	US-10-424-599-221580	Sequence 221580,
36	452.5	20.7	325	12	US-10-425-114-50591	Sequence 50591, A
37	447	20.5	322	14	US-10-162-012-42	Sequence 42, Appl
38	447	20.5	322	14	US-10-176-306-51	Sequence 51, Appl
39	447	20.5	322	14	US-10-176-306-57	Sequence 57, Appl
40	447	20.5	322	15	US-10-162-102-42	Sequence 42, Appl
41	435	19.9	356	12	US-10-424-599-221581	Sequence 221581,
42	432.5	19.8	289	14	US-10-176-306-58	Sequence 58, Appl
43	384.5	17.6	287	12	US-10-424-599-185725	Sequence 185725,
44	384.5	17.6	371	14	US-10-156-761-14576	Sequence 14576, A
45	356	16.3	318	9	US-09-738-626-4910	Sequence 4910, Ap

ALIGNMENTS

RESULT 1
US-10-024-652-2570
; Sequence 2570, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2570
; LENGTH: 429
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-2570

Query Match 100.0%; Score 2183; DB 15; Length 429;
Best Local Similarity 100.0%; Pred. No. 3.3e-218;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
Db 1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

```
Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

RESULT 2
US-10-024-652-2572
; Sequence 2572, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2572
; LENGTH: 429
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-2572

Query Match 100.0%; Score 2183; DB 15; Length 429;
Best Local Similarity 100.0%; Pred. No. 3.3e-218;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

RESULT 3
US-10-024-652-2575
; Sequence 2575, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2575
; LENGTH: 429
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-2575

Query Match 100.0%; Score 2183; DB 15; Length 429;
Best Local Similarity 100.0%; Pred. No. 3.3e-218;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

RESULT 4
US-10-024-652-2576
; Sequence 2576, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2576
; LENGTH: 429
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-2576

Query Match 100.0%; Score 2183; DB 15; Length 429;
Best Local Similarity 100.0%; Pred. No. 3.3e-218;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

RESULT 5
US-10-024-652-2577
; Sequence 2577, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2577
; LENGTH: 429
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-2577

Query Match 100.0%; Score 2183; DB 15; Length 429;
Best Local Similarity 100.0%; Pred. No. 3.3e-218;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

```
Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 6
US-10-024-652-2578
; Sequence 2578, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc
;  APPLICANT:  Challita-Eid, Pia M.
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Afar, Daniel E.H.
;  APPLICANT:  Hubert, Rene S.
;  APPLICANT:  Mitchell, Steve Chappell
;  APPLICANT:  Levin, Elana
;  APPLICANT:  Morrison, Karen Jane Meyrick
;  APPLICANT:  Raitano, Arthur B.
;  APPLICANT:  Jakobovits, Aya
;  TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
;  TITLE OF INVENTION:  Transporter Protein Entitled 108P5H8 Useful in Treatment and
;  TITLE OF INVENTION:  Detection of Cancer
;  FILE REFERENCE: 51158-20025.00
;  CURRENT APPLICATION NUMBER: US/10/024,652
;  CURRENT FILING DATE:  2002-06-28
;  PRIOR APPLICATION NUMBER: 60/256,210
;  PRIOR FILING DATE: 2000-12-15
;  NUMBER OF SEQ ID NOS: 2598
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2578
;   LENGTH: 429
;   TYPE: PRT
;   ORGANISM: homo sapien
US-10-024-652-2578

  Query Match            100.0%;  Score 2183;  DB 15;  Length 429;
  Best Local Similarity  100.0%;  Pred. No. 3.3e-218;
  Matches  429;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 7
US-10-295-027-544
; Sequence 544, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
;  APPLICANT: Afar, Daniel
;  APPLICANT:  Aziz, Natasha
;  APPLICANT:  Ginsberg, Wendy M.
;  APPLICANT:  Gish, Kurt C.
;  APPLICANT:  Glynne, Richard
;  APPLICANT:  Hevezi, Peter A.
;  APPLICANT:  Mack, David H.
;  APPLICANT:  Murray, Richard
;  APPLICANT:  Watson, Susan R.
;  APPLICANT:  Eos Biotechnology, Inc.
;  TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;  TITLE OF INVENTION:  Methods of Screening for Modulators of Cancer
;  FILE REFERENCE: 018501-012500US
;  CURRENT APPLICATION NUMBER: US/10/295,027
;  CURRENT FILING DATE:  2002-11-13
;  PRIOR APPLICATION NUMBER: US 09/663,733
;  PRIOR FILING DATE: 2000-09-15
;  PRIOR APPLICATION NUMBER: US 60/350,666
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/335,394
;  PRIOR FILING DATE: 2001-11-15
;  PRIOR APPLICATION NUMBER: US 60/332,464
;  PRIOR FILING DATE: 2001-11-21
;  PRIOR APPLICATION NUMBER: US 60/334,393
;  PRIOR FILING DATE: 2001-11-29
;  PRIOR APPLICATION NUMBER: US 60/340,376
;  PRIOR FILING DATE: 2001-12-14
;  PRIOR APPLICATION NUMBER: US 60/347,211
;  PRIOR FILING DATE: 2002-01-08
;  PRIOR APPLICATION NUMBER: US 60/347,349
;  PRIOR FILING DATE: 2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/355,250
;  PRIOR FILING DATE: 2002-02-08
;  PRIOR APPLICATION NUMBER: US 60/356,714
;  PRIOR FILING DATE: 2002-02-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 1386
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 544
;   LENGTH: 429
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-295-027-544

  Query Match            100.0%;  Score 2183;  DB 15;  Length 429;
  Best Local Similarity  100.0%;  Pred. No. 3.3e-218;
  Matches  429;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
```

```
Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 8
US-09-957-708-27
; Sequence 27, Application US/09957708
; Publication No. US20030031678A1
; GENERAL INFORMATION:
;  APPLICANT: Sun, Yongming
;  APPLICANT:  Recipon, Herve
;  APPLICANT:  Cafferkey, Robert
;  APPLICANT:  Ali, Shujath
;  TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
;  TITLE OF INVENTION:  Genes
;  FILE REFERENCE: DEX-0239
;  CURRENT APPLICATION NUMBER: US/09/957,708
;  CURRENT FILING DATE:  2001-09-19
;  PRIOR APPLICATION NUMBER: 60/233,746
;  PRIOR FILING DATE: 2000-09-19
;  NUMBER OF SEQ ID NOS: 40
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 438
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-957-708-27

  Query Match           100.0%;  Score 2183;  DB 10;  Length 438;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-218;
  Matches  429;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 9
US-10-345-680-53
; Sequence 53, Application US/10345680
; Publication No. US20030148394A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals, Inc.
;  APPLICANT:  Silos-Santiago, Inmaculada
;  APPLICANT:  Venkateswarlu, Karicheti
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  UROLOGICAL DISORDERS USING 1435, 559, 34021, 44099, 25278,
;  TITLE OF INVENTION:  641, 260, 55089, 21407, 42032, 46656, 62553, 302, 323,
;  TITLE OF INVENTION:  12303, 985, 13237, 13601, 18926, 318, 2058 OR 6351 MOLECULES.
;  FILE REFERENCE: MPI02-012P1RNM_OMNI
;  CURRENT APPLICATION NUMBER: US/10/345,680
;  CURRENT FILING DATE:  2003-01-16
;  PRIOR APPLICATION NUMBER: US 60/349,511
;  PRIOR FILING DATE: 2002-01-18
;  PRIOR APPLICATION NUMBER: US 60/360,500
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/365,041
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/374,063
;  PRIOR FILING DATE: 2002-04-19
;  PRIOR APPLICATION NUMBER: US 60/403,468
;  PRIOR FILING DATE: 2002-08-14
;  PRIOR APPLICATION NUMBER: US 60/414,262
;  PRIOR FILING DATE: 2002-09-27
;  PRIOR APPLICATION NUMBER: US 60/419,986
;  PRIOR FILING DATE: 2002-10-21
;  PRIOR APPLICATION NUMBER: US 60/423,809
;  PRIOR FILING DATE: 2002-11-05
;  PRIOR APPLICATION NUMBER: US 60/429,797
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 66
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53
;   LENGTH: 429
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-345-680-53

  Query Match           99.8%;  Score 2179;  DB 14;  Length 429;
  Best Local Similarity  99.8%;  Pred. No. 8.7e-218;
  Matches  428;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFEFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 10
US-10-024-652-2574
; Sequence 2574, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc
;  APPLICANT:  Challita-Eid, Pia M.
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Afar, Daniel E.H.
;  APPLICANT:  Hubert, Rene S.
;  APPLICANT:  Mitchell, Steve Chappell
;  APPLICANT:  Levin, Elana
;  APPLICANT:  Morrison, Karen Jane Meyrick
;  APPLICANT:  Raitano, Arthur B.
;  APPLICANT:  Jakobovits, Aya
;  TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
;  TITLE OF INVENTION:  Transporter Protein Entitled 108P5H8 Useful in Treatment and
;  TITLE OF INVENTION:  Detection of Cancer
;  FILE REFERENCE: 51158-20025.00
;  CURRENT APPLICATION NUMBER: US/10/024,652
;  CURRENT FILING DATE:  2002-06-28
;  PRIOR APPLICATION NUMBER: 60/256,210
;  PRIOR FILING DATE: 2000-12-15
;  NUMBER OF SEQ ID NOS: 2598
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2574
;   LENGTH: 429
;   TYPE: PRT
;   ORGANISM: homo sapien
US-10-024-652-2574

  Query Match             99.8%;  Score 2179;  DB 15;  Length 429;
  Best Local Similarity   99.8%;  Pred. No. 8.7e-218;
  Matches  428;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFEFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 11
US-10-024-652-2579
; Sequence 2579, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc
;  APPLICANT:  Challita-Eid, Pia M.
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Afar, Daniel E.H.
;  APPLICANT:  Hubert, Rene S.
;  APPLICANT:  Mitchell, Steve Chappell
;  APPLICANT:  Levin, Elana
;  APPLICANT:  Morrison, Karen Jane Meyrick
;  APPLICANT:  Raitano, Arthur B.
;  APPLICANT:  Jakobovits, Aya
;  TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
;  TITLE OF INVENTION:  Transporter Protein Entitled 108P5H8 Useful in Treatment and
;  TITLE OF INVENTION:  Detection of Cancer
;  FILE REFERENCE: 51158-20025.00
;  CURRENT APPLICATION NUMBER: US/10/024,652
;  CURRENT FILING DATE:  2002-06-28
;  PRIOR APPLICATION NUMBER: 60/256,210
;  PRIOR FILING DATE: 2000-12-15
;  NUMBER OF SEQ ID NOS: 2598
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2579
;   LENGTH: 429
;   TYPE: PRT
;   ORGANISM: homo sapien
US-10-024-652-2579

  Query Match             99.8%;  Score 2179;  DB 15;  Length 429;
  Best Local Similarity   99.8%;  Pred. No. 8.7e-218;
  Matches  428;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFEFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

```
RESULT 12
US-10-024-652-2580
; Sequence 2580, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc
;  APPLICANT:  Challita-Eid, Pia M.
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Afar, Daniel E.H.
;  APPLICANT:  Hubert, Rene S.
;  APPLICANT:  Mitchell, Steve Chappell
;  APPLICANT:  Levin, Elana
;  APPLICANT:  Morrison, Karen Jane Meyrick
;  APPLICANT:  Raitano, Arthur B.
;  APPLICANT:  Jakobovits, Aya
;  TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
;  TITLE OF INVENTION:  Transporter Protein Entitled 108P5H8 Useful in Treatment and
;  TITLE OF INVENTION:  Detection of Cancer
;  FILE REFERENCE: 51158-20025.00
;  CURRENT APPLICATION NUMBER: US/10/024,652
;  CURRENT FILING DATE:  2002-06-28
;  PRIOR APPLICATION NUMBER: 60/256,210
;  PRIOR FILING DATE: 2000-12-15
;  NUMBER OF SEQ ID NOS: 2598
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2580
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: homo sapien
US-10-024-652-2580

  Query Match             99.7%;  Score 2176;  DB 15;  Length 428;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-217;
  Matches  428;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSS 428
              ||||||||
Db        421 CANCQSSS 428

RESULT 13
US-10-012-697-1545
; Sequence 1545, Application US/10012697
; Publication No. US20030215803A1
; GENERAL INFORMATION:
;  APPLICANT: Escobedo, Jaime
;  APPLICANT:  Garcia, Pablo Dominguez
;  APPLICANT:  Kassam, Altaf
;  APPLICANT:  Lamson, George
;  APPLICANT:  Scott, Beth
;  APPLICANT:  Drmanac, Radoje
;  APPLICANT:  Crkvenjakov, Radomir
;  APPLICANT:  Dickson, Mark
;  APPLICANT:  Drmanac, Snezana
;  APPLICANT:  Labat, Ivan
;  APPLICANT:  Leshkowitz, Dena
;  APPLICANT:  Kita, David
;  APPLICANT:  Garcia, Veronica
;  APPLICANT:  Jones, Lee William
;  APPLICANT:  Stache-Crain, Birgit
;  TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS
;  TITLE OF INVENTION:  ISOLATED FROM HUMAN PROSTATE
;  FILE REFERENCE: 2300-16252
;  CURRENT APPLICATION NUMBER: US/10/012,697
;  CURRENT FILING DATE:  2003-01-21
;  PRIOR APPLICATION NUMBER: 60/254,648
;  PRIOR FILING DATE: 2000-12-07
;  PRIOR APPLICATION NUMBER: 60/275,668
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 1568
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1545
;   LENGTH: 430
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 181
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 181
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 181
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 181
;   OTHER INFORMATION: Xaa = Any Amino Acid
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: 181
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-10-012-697-1545

  Query Match             99.2%;  Score 2166.5;  DB 15;  Length 430;
  Best Local Similarity   99.5%;  Pred. No. 1.8e-216;
  Matches  428;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRL- 179
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLG 180

Qy        180 EVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQS 239
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 XVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQS 240
```

```
Qy        240 GHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKP 299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKP 300

Qy        300 EYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVED 359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVED 360

Qy        360 LNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDR 419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDR 420

Qy        420 TCANCQSSSP 429
              ||||||||||
Db        421 TCANCQSSSP 430
```

RESULT 14
US-10-024-652-2581
; Sequence 2581, Application US/10024652
; Publication No. US20030219738A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Faris, Mary
; APPLICANT: Afar, Daniel E.H.
; APPLICANT: Hubert, Rene S.
; APPLICANT: Mitchell, Steve Chappell
; APPLICANT: Levin, Elana
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Encoded Zinc
; TITLE OF INVENTION: Transporter Protein Entitled 108P5H8 Useful in Treatment and
; TITLE OF INVENTION: Detection of Cancer
; FILE REFERENCE: 51158-20025.00
; CURRENT APPLICATION NUMBER: US/10/024,652
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: 60/256,210
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 2598
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2581
; LENGTH: 429
; TYPE: PRT
; ORGANISM: homo sapien
US-10-024-652-2581

Query Match 91.2%; Score 1990.5; DB 15; Length 429;
Best Local Similarity 90.7%; Pred. No. 3.8e-198;
Matches 390; Conservative 19; Mismatches 18; Indels 3; Gaps 2;

```
Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||| |||||||:|||||||||||||||||| ||  |||||||||||||||| ||||||
Db          1 MAGPGAWKRLKSLLRKDDAPLFLNDTSAFDFLDEVSDEGLSRFNKLRVVVADDDSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||| |||||||||||:||||||||||:| ||||||:||:||||||| ||||||||
Db         61 PVNGAHPALQADDDSLLDQELPLTNSQLSLKMDPCDNCSKRRELLKQRKVKTRLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              |||||||||||||:|||||||||||||||||||||||||||||||||:||||||||||
Db        121 YLLFMIGELVGGYMANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTRRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||:|||:||||||||:|||||||||||||:|||||||||||||||||||||||
Db        181 VLSAMISVMLVYVLMGFLLYEAMQRTIHMNYEINGDVMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHS--HSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFK 298
              | ||  ||||||||||:  |    :||||||||||||||||||||||||||||||||||
Db        241 HHHSHAHSHSLPSNSPSMVSS-GHSHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFK 299

Qy        299 PEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVE 358
              ||||||||||||:||||||||| |||||||||||||||||||||||||:||||||||||
Db        300 PEYKIADPICTYIFSLLVAFTTLRIIWDTVVIILEGVPSHLNVDYIKESLMKIEDVYSVE 359

Qy        359 DLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVD 418
              ||||||||||:|||||:|||||||||||||||| ||||||||||:||:||||||||
Db        360 DLNIWSLTSGKATAIVHMQLIPGSSSKWEEVQSKAKHLLLNTFGMYKCTVQLQSYRQEAT 419

Qy        419 RTCANCQSSS 428
              ||||||||||
Db        420 RTCANCQSSS 429
```

RESULT 15
US-09-954-342-37
; Sequence 37, Application US/09954342
; Publication No. US20030170838A1
; GENERAL INFORMATION:
; APPLICANT: MISHRA, VISHNU S.
; APPLICANT: SPYTEK, KIMBERLY ANN
; APPLICANT: TAUPIER, RAYMOND J.
; APPLICANT: VERNET, CORINE A.
; APPLICANT: COLMAN, STEVEN D.
; APPLICANT: GORMAN, LINDA
; APPLICANT: TCHERNEV, VELIZAR T.
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: SHENOY, SURESH
; APPLICANT: PADIGARU, MURALIDHARA
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENNDA
; APPLICANT: MILLET, ISABELLE
; APPLICANT: PEYMAN, JOHN
; APPLICANT: STONE, DAVID
; APPLICANT: GUNTHER, ERIK
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: LI, LI
; APPLICANT: RASTELLI, LUCA
; APPLICANT: ZERHUSEN, BRYAN
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-114
; CURRENT APPLICATION NUMBER: US/09/954,342
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/233,382
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 60/240,498
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/260,284
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/260,973
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/264,794
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/238,398
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/232,675
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/274,862
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/233,801
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/232,676
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/233,960
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,402
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 60/233,521
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/233,522
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/232,679
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 104

```
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Rattus norvegicus
US-09-954-342-37

  Query Match             34.2%;  Score 747;  DB 10;  Length 359;
  Best Local Similarity   42.0%;  Pred. No. 1.1e-68;
  Matches  148;  Conservative   78;  Mismatches  108;  Indels   18;  Gaps    4;

Qy          82 PLTNSQLSLKVDSCDNCSKQREI-----LKQRKVKARLTIAAVLYLLFMIGELVGGYIAN 136
               |:    : | | |   |  |::       :::: : :| :|: : |:|||||::|||:|
Db          20 PVNLPSVELAVQSNHYCHAQKDSGSHPNSEKQRARRKLYVASAICLVFMIGEIIGGYLAQ 79

Qy         137 SLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMG 196
               |||||||| |:||| :::::::| :||:||:  ||   ||: | |:| |::||| :::: |
Db          80 SLAIMTDAAHLLTDFASMLISLFSLWVSSRPATKTMNFGWQRAEILGALLSVLSIWVVTG 139

Qy         197 FLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTR 256
                |:| |||| |  :||| || ||||:   ||||:|||  |:|||| ||| ||   :|
Db         140 VLVYLAVQRLISGDYEIKGDTMLITSGCAVAVNIIMGLALHQSGHGHSHGHSHEDSS--- 196

Qy         257 GSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLV 316
                     | : :|||||:| :|||:|||||||:|||||  ||||||  |||||::||:||
Db         197 -------QQQQNPSVRAAFIHVVGDLLQSVGVLVAAYIIYFKPEYKYVDPICTFLFSILV 249

Qy         317 AFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHI 376
               ||  |: | :::::|| |  ::   :|  |: :: | ::  |:||:||  :    |||
Db         250 LGTTLTILRDVILVLMEGTPKGVDFTTVKNLLLSVDGVEALHSLHIWALTVAQPVLSVHI 309

Qy         377 QLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQSSS 428
                :    : | |   |   |   |  :  |||::|| ::: ::|  ||  |
Db         310 AIAQNVDA--QAVLKVARDRLQGKFNFHTMTIQIESYSEDM-KSCQECQGPS 358


Search completed: July 29, 2004, 17:46:17
Job time : 48 secs
```

Blank sheet

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 29, 2004, 17:40:40 ; Search time 39 Seconds
(without alignments)
3470.700 Million cell updates/sec

Title: US-10-024-652-2570
Perfect score: 2183
Sequence: 1 MAGSGAWKRLKSMLRKDDAP.....LQSYRQEVDRTCANCQSSSP 429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2183	100.0	429	4	Q8TC39	Q8tc39 homo sapien
2	742	34.0	146	4	Q96JT7	Q96jt7 homo sapien
3	652.5	29.9	472	5	Q9V471	Q9v471 drosophila
4	648.5	29.7	369	4	Q8IWU4	Q8iwu4 homo sapien
5	643.5	29.5	367	11	Q8BGG0	Q8bgg0 mus musculu
6	638.5	29.2	382	5	Q22541	Q22541 caenorhabdi
7	638	29.2	320	4	Q8TCL3	Q8tcl3 homo sapien
8	627	28.7	440	5	Q8IP48	Q8ip48 drosophila
9	615	28.2	388	4	Q8TC03	Q8tc03 homo sapien
10	607.5	27.8	410	5	O45923	O45923 caenorhabdi
11	603.5	27.6	391	5	O45922	O45922 caenorhabdi
12	599.5	27.5	669	5	Q9VKA3	Q9vka3 drosophila
13	565.5	25.9	498	5	Q8T0G1	Q8t0g1 drosophila
14	555	25.4	398	10	O81036	O81036 arabidopsis
15	547	25.1	398	10	Q9ZT63	Q9zt63 arabidopsis
16	546	25.0	392	10	Q94B00	Q94b00 thlaspi goe
17	543.5	24.9	396	10	Q93XE9	Q93xe9 thlaspi cae
18	539	24.7	323	4	Q9BRI3	Q9bri3 homo sapien
19	537.5	24.6	418	10	Q84NH0	Q84nh0 oryza sativ
20	536.5	24.6	421	10	Q947R8	Q947r8 eucalyptus
21	534	24.5	385	10	Q9ZW23	Q9zw23 arabidopsis
22	532	24.4	378	10	Q9LXS1	Q9lxs1 arabidopsis
23	530	24.3	360	10	Q94AZ9	Q94az9 thlaspi goe
24	526.5	24.1	334	10	Q9M271	Q9m271 arabidopsis
25	469.5	21.5	334	5	Q8SQT3	Q8sqt3 encephalito
26	453	20.8	299	16	Q81F93	Q81f93 bacillus ce
27	446	20.4	299	16	Q81SA9	Q81sa9 bacillus an
28	444.5	20.4	300	16	Q8ERA7	Q8era7 oceanobacil
29	440.5	20.2	319	16	Q8NYN6	Q8nyn6 staphylococ
30	437.5	20.0	319	16	Q99X53	Q99x53 staphylococ
31	429	19.7	556	5	Q8IBU1	Q8ibu1 plasmodium
32	418.5	19.2	313	16	Q8CPX7	Q8cpx7 staphylococ
33	407	18.6	308	16	O67168	O67168 aquifex aeo
34	399.5	18.3	317	2	O05214	O05214 bacillus st
35	396.5	18.2	316	2	Q9R494	Q9r494 bacillus st
36	391	17.9	303	16	Q8Y480	Q8y480 listeria mo
37	387	17.7	303	16	Q927R8	Q927r8 listeria in
38	385	17.6	313	16	Q83SA2	Q83sa2 shigella fl
39	384.5	17.6	371	16	Q826X7	Q826x7 streptomyce
40	383.5	17.6	317	16	Q8F433	Q8f433 leptospira
41	380.5	17.4	301	16	Q88RV3	Q88rv3 pseudomonas
42	377	17.3	297	16	Q83EH0	Q83eh0 coxiella bu
43	376.5	17.2	449	5	Q9VZR4	Q9vzr4 drosophila
44	372.5	17.1	323	16	Q8CNH2	Q8cnh2 staphylococ
45	370	16.9	311	16	O07084	O07084 bacillus su

ALIGNMENTS

RESULT 1
Q8TC39
ID Q8TC39 PRELIMINARY; PRT; 429 AA.
AC Q8TC39;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Solute carrier family 30 (Zinc transporter), member 4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC026089; AAH26089.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008324; F:cation transporter activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR InterPro; IPR002524; Cation_efflux.
DR Pfam; PF01545; Cation_efflux; 1.
DR TIGRFAMs; TIGR01297; CDF; 1.
SQ SEQUENCE 429 AA; 47482 MW; 97B7FCBE881C8C32 CRC64;

Query Match 100.0%; Score 2183; DB 4; Length 429;
Best Local Similarity 100.0%; Pred. No. 5.6e-183;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
                 |||||||||
Db        421 CANCQSSSP 429


RESULT 2
Q96JT7
ID   Q96JT7      PRELIMINARY;      PRT;   146 AA.
AC   Q96JT7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative zinc transporter 4 (Fragment).
GN   ZNT4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=malignant prostate;
RA   Nong G., Buszko M.L., Sweeney C.A., Furman J., Rice L.P.;
RT   "Expression of zinc transporters in human prostate.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY029482; AAK40257.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
FT   NON_TER       1       1
FT   NON_TER     146     146
SQ   SEQUENCE   146 AA;  16099 MW;  9831FB5FE9CFBE03 CRC64;

  Query Match             34.0%;  Score 742;  DB 4;  Length 146;
  Best Local Similarity   99.3%;  Pred. No. 4e-57;
  Matches  144;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        242 RHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEY 301
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 RHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEY 60

Qy        302 KIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLN 361
                 ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KIAXPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLN 120

Qy        362 IWSLTSGKSTAIVHIQLIPGSSSKW 386
                 |||||||||||||||||||||||||
Db        121 IWSLTSGKSTAIVHIQLIPGSSSKW 145


RESULT 3
Q9V471
ID   Q9V471      PRELIMINARY;      PRT;   472 AA.
AC   Q9V471;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   BG:DS07295.1 protein (RE54080p).
GN   BG:DS07295.1 OR CG3994.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=99403001; PubMed=10471707;
RA   Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA   Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA   Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA   Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA   Celniker S., Rubin G.M.;
RT   "An exploration of the sequence of a 2.9-Mb region of the genome of
RT   Drosophila melanogaster: the Adh region.";
RL   Genetics 153:179-219(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA   Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA   Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA   Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA   Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
```

```
RA   Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA   Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA   Zieran L.L., Rubin G.M.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003646; AAF53443.1; -.
DR   EMBL; AE003412; AAF44926.1; -.
DR   EMBL; AY071460; AAL49082.1; -.
DR   FlyBase; FBgn0028516; BG:DS07295.1.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   472 AA;  51036 MW;  077037ACFAAE3BF5 CRC64;

  Query Match             29.9%;  Score 652.5;  DB 5;  Length 472;
  Best Local Similarity   31.1%;  Pred. No. 1.4e-48;
  Matches  155;  Conservative   91;  Mismatches  135;  Indels  117;  Gaps   12;

Qy         13 MLRKDDAPLFLNDTSAFDFSD----------EAGDEGLSRFNKLRVVVADDGSEAPERPV 62
              | | :| |:    |:      |:          | |  |::       ||| :|
Db          1 MSRNEDTPIAKRDSGRTRRSNYGTAPSFHLMEQGQPGVN-------VVAGNG-------- 45

Qy         63 NGAHPTLQADDDSLLDQDLPLTNSQL------SLKVDSCDNCSKQREILKQRKVKARLTI 116
              | ||             | | :|:      |  |:  |:| : |      | : :| |
Db         46 NNNHPA------------TPATPAQIFCLHGRSNNVEVRDHCHRARSEGVDVKARRKLII 93

Qy        117 AAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGF 176
              |::| |:||| |:||| ::||||| ||| |:||| :: :::| |:|:: :  |:| :||:
Db         94 ASILCLVFMIAEIVGGVLSNSLAIATDAAHLLTDFASFMISLFAIWIAGRPSTQRMSFGW 153

Qy        177 HRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLL 236
              :| ||: || || ::::: | |:: |: | |  :||:|  |||||: : : ||||||  |
Db        154 YRAEVIGAMASVFMIWVITGILVWLAIGRLISGDYEVNAKIMLITSGLAILVNVIMGVQL 213

Qy        237 NQ-------SGHRHSH------SHSLPSNSPTRGSGCER---------------------- 262
                      || |||       ||   :::|   |  :|
Db        214 QHGHSHGLGGGHGHSHGGSKNASHVQATSTPCSDSPSQRIEGGVAYAPEDAELPGGGLPT 273

Qy        263 ----------------------------NHGQD------SLAVRAAFVHALGDLVQSVG 287
                                         :||        :: |||| :| :||::||||
Db        274 FSYQNTKLVDPTLDLEIAAVLAETAPGSHHHGGPVGREAVNMNVRAALIHVIGDVIQSVG 333

Qy        288 VLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEA 347
              | :|| :| | ||| | |||||:|||::| |||| |: | :::::|| |::::   : :
Db        334 VFVAAGVIYFWPEYSIVDPICTFVFSIIVLFTTFTIMKDALLVLMEGTPNYMHYAEVLQI 393

Qy        348 LMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSK-WEEVQSKANHLLLNTFGMYRC 406
                 || |  | :| ||:|:  |      |: :   :: |   :  : | ||  | |
Db        394 FQGIEGVERVHNLRIWALSINKVALSAHLAIAENANPKRILDAATSAVHLRYNFF---ET 450

Qy        407 TIQLQSYRQEVDRTCANC 424
              |||:: |  ::: :|  |
Db        451 TIQIEDYTAQME-SCLQC 467


RESULT 4
Q8IWU4
ID   Q8IWU4      PRELIMINARY;      PRT;   369 AA.
AC   Q8IWU4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Zinc transporter ZnT-8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RA   Seve M., Devergnas S., Chimienti F., Chantegrel J., Favier A.;
RT   "ZnT-8, a pancreatic specific zinc transporter.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY117411; AAM80562.1; -.
DR   Genew; HGNC:20303; SLC30A8.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   369 AA;  40728 MW;  A568300E83DCBB22 CRC64;

  Query Match             29.7%;  Score 648.5;  DB 4;  Length 369;
  Best Local Similarity   38.0%;  Pred. No. 2.3e-48;
  Matches  139;  Conservative   75;  Mismatches  107;  Indels   45;  Gaps    6;

Qy         83 LTNSQLSLKVDSCDNCSKQR-EILKQ-------------------RKVKARLTIAAVLY 121
              | : :|  |  : | | ::| | |:                      | :|  |: :
Db         22 LESVELQQKPVNKDQCPRERPEELESGGMYHCHSGSKPTEKGANEYAYAKWKLCSASAIC 81

Qy        122 LLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEV 181
              :||| |:|||:|| |||::||| |:| ||:: :|:| :|||||| |:|| |||:|| |:
Db         82 FIFMIAEVVGGHIAGSLAVVTDAAHLLIDLTSFLLSLFSLWLSSKPPSKRLTFGWHRAEI 141

Qy        182 LSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQS-- 239
              | |::|:| :::: | |:| | :| :: :|:|   :|:| ::  || |:::  :|:|
Db        142 LGALLSILCIWVVTGVLVYLACERLLYPDYQIQATVMIIVSSCAVAANIVLTVVLHQRCL 201

Qy        240 GHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKP 299
              || |                    | : :||||||||||| ||: |||:| || |||
Db        202 GHNHKEV------------------QANASVRAAFVHALGDLFQSISVLISALIIYFKP 242

Qy        300 EYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVED 359
              |||||||||::||:||  :|  |: |  ::::|||  ||   :|| :: :: | ||
Db        243 EYKIADPICTFIFSILVLASTITILKDFSILLMEGVPKSLNYSGVKELILAVDGVLSVHS 302

Qy        360 LNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDR 419
              |:|||||  :    |:      |  : |: :   |  :| |:  |||::|   : |
Db        303 LHIWSLTMNQVILSAHVATAASRDS--QVVRREIAKALSKSFTMHSLTIQMES-PVDQDP 359

Qy        420 TCANCQ 425
              |  |:
Db        360 DCLFCE 365


RESULT 5
Q8BGG0
ID   Q8BGG0      PRELIMINARY;      PRT;   367 AA.
AC   Q8BGG0;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Weakly similar to zinc transporter 2.
GN   C820002P14RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Pancreas;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
```

```
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK050482; BAC34281.1; -.
DR   EMBL; AK050494; BAC34289.1; -.
DR   MGD; MGI:2442682; C820002P14Rik.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   367 AA;  40222 MW;  9C9484A561DE58E4 CRC64;

  Query Match             29.5%;  Score 643.5;  DB 11;  Length 367;
  Best Local Similarity   38.8%;  Pred. No. 6.2e-48;
  Matches  130;  Conservative   81;  Mismatches   97;  Indels   27;  Gaps    5;

Qy         95 CDNCSKQ--REILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLS 152
              | |:|      ||   | ||   |: :  :||: |:|||::| ||||:||| |:| ||:
Db         52 CHNSAKATGNRSSKQAHAKWRLCAASAICFIFMVAEVVGGHVAGSLAILTDAAHLLIDLT 111

Qy        153 AIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYE 212
              : :|:| :|||||: |:|| |||::| |:| |::||| :::: | ||| | :| :: :|:
Db        112 SFLLSLFSLWLSSRPPSKRLTFGWYRAEILGALLSVLCIWVVTGVLLYLACERLLYPDYQ 171

Qy        213 INGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVR 272
              |  ||:  :   || |:::  :|:|    ::|                      | : :||
Db        172 IQAGIMITVSGCAVAANIVLTMILHQRNFGYNHKDV----------------QANASVR 214

Qy        273 AAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIIL 332
              |||||||||: ||: |||:| || |||:||||||:||::||:|  :|  |: |  :::: 
Db        215 AAFVHALGDVFQSISVLISALIIYFKPDYKIADPVCTFIFSILVLASTVMILKDFSILLM 274

Qy        333 EGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSK 392
              ||||  |: : :|| :: :: | ||  |:|||||  :    ||:     :::  :: ||
Db        275 EGVPKGLSYNSVKEIILAVDGVISVHSLHIWSLTVNQVILSVHV-----ATAASQDSQSV 329

Qy        393 ANHL--LLNTFGMYRCTIQLQSYRQEVDRTCANCQ 425
                 :   |::| ::  |||::|   : | :|  |:
Db        330 RTGIAQALSSFDLHSLTIQIESAADQ-DPSCLLCE 363


RESULT 6
Q22541
ID   Q22541      PRELIMINARY;      PRT;   382 AA.
AC   Q22541;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   T18D3.3 protein.
GN   T18D3.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McMurray A.A.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; Z68119; CAA92193.1; -.
DR   PIR; T24963; T24963.
DR   WormPep; T18D3.3; CE03664.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   382 AA;  42237 MW;  44A02AEF60B89484 CRC64;

  Query Match             29.2%;  Score 638.5;  DB 5;  Length 382;
  Best Local Similarity   39.8%;  Pred. No. 1.8e-47;
  Matches  136;  Conservative   70;  Mismatches  105;  Indels   31;  Gaps    8;

Qy         91 KVDSCDNCSKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTD 150
              : || |:          |:    | :  || | ||: |::|| :| ||||:||| |:||| 
Db         59 EADSTDSHD------SNRRATRILWLTVVLCLFFMVCEVIGGVLAGSLAIVTDAAHLLTD 112

Qy        151 LSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMN 210
              ::::::| :|::: : |::: :|||||  ||| |  || |::|: | |:  |: | :  :
Db        113 FASVLISLFSLYIARRPPSQKMSFGFHRAEVLGAFFSVFLIWIVTGVLVVLAIMRIVSGD 172

Qy        211 YEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQ---- 266
              ||: | || :|||:|| ||::|  ||   || |||           | |    |
Db        173 YEVEGGIMALTAALGVVVNLVMLALLYFGGHSHSH----------GGGSSHGHSHGGGN 221

Qy        267 -DSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIW 325
               |:: ||||:| ||||:||:|||:||  | |:| : | |||| |||::|  ||  |:
Db        222 GDNINVRAAFIHVLGDLLQSLGVLVAALFIYFQPSWVIIDPICTLVFSVIVLCTTIYILR 281

Qy        326 DTVVIILEGVPSHLNVDYIK--EALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSS 383
              | :::||| ||  |:|: |   :|  || |  | || |||||  |    ||:::  |
Db        282 DAMIVLLEGRPS--NIDFAKVFSSLEDIEGVKKVHDLRIWSLTMDKIALSVHLEIDANSQ 339

Qy        384 SKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT-CANC 424
              |  : :  :  :|  |: ::  |||:: :    :|: |  |
Db        340 S--QSILRETRKMLKQTYNVHEITIQIEEF--GANRSDCGKC 377


RESULT 7
Q8TCL3
ID   Q8TCL3      PRELIMINARY;      PRT;   320 AA.
AC   Q8TCL3;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   DKFZP564F1062.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Blum H., Bauersachs S., Mewes H.W., Gassenhuber J., Wiemann S.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL713790; CAD28545.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   320 AA;  35053 MW;  5D7325EBD1758EC9 CRC64;

  Query Match             29.2%;  Score 638;  DB 4;  Length 320;
  Best Local Similarity   41.0%;  Pred. No. 1.6e-47;
  Matches  130;  Conservative   69;  Mismatches   94;  Indels   24;  Gaps    4;

Qy        111 KARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTK 170
              | :|  |: :  :||| |:|||:|| |||::||| |:| ||:: :|:| :|||||| |:|
Db         22 KWKLCSASAICFIFMIAEVVGGHIAGSLAVVTDAAHLLIDLTSFLLSLFSLWLSSKPPSK 81

Qy        171 RFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNV 230
              | |||:|| |:| |::|:| :::: | |:| | :| :: :|:|   :|:| ::  || |:
```

```
Db         82 RLTFGWHRAEILGALLSILCIWVVTGVLVYLACERLLYPDYQIQATVMIIVSSCAVAANI 141

Qy        231 IMGFLLNQS--GHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGV 288
              ::  :|:|   || |                        | : :|||||||| ||| ||: |
Db        142 VLTVVLHQRCLGHNHKEV------------------QANASVRAAFVHAPGDLFQSISV 182

Qy        289 LIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEAL 348
              ||:| || ||||||||||||||::||:||  :|  |: |  ::::|||| ||   :|| :
Db        183 LISALIIYFKPEYKIADPICTFIFSILVLASTITILKDFSILLMEGVPKSLNYSGVKELI 242

Qy        349 MKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTI 408
              : :: | ||  |:|||||  :     |:       | : |: :    |  :| |:  ||
Db        243 LAVDGVLSVHCLHIWSLTMNQVILSAHVATAASRDS--QVVRREIAKALSKSFTMHSLTI 300

Qy        409 QLQSYRQEVDRTCANCQ 425
              |::|   : |  |  |:
Db        301 QMES-PVDQDPDCLFCE 316


RESULT 8
Q8IP48
ID   Q8IP48      PRELIMINARY;      PRT;   440 AA.
AC   Q8IP48;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CG3994-PB.
GN   BG:DS07295.1 OR CG3994.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003646; AAN10893.1; -.
DR   FlyBase; FBgn0028516; BG:DS07295.1.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   440 AA;  47993 MW;  82E291198572C76C CRC64;

  Query Match           28.7%;  Score 627;  DB 5;  Length 440;
  Best Local Similarity  33.6%;  Pred. No. 2.2e-46;
  Matches  134;  Conservative   81;  Mismatches  110;  Indels   74;  Gaps    7;

Qy         96 DNCSKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAII 155
              |:| : |      | : :| ||::| |:||| |:||| ::||||| ||| |:||| :: :
Db         41 DHCHRARSEGVDVKARRKLIIASILCLVFMIAEIVGGVLSNSLAIATDAAHLLTDFASFM 100

Qy        156 LTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEING 215
              ::| |:|:: :  |:| :||::| ||: || || ::::: | |:: |: | |  :||:|
Db        101 ISLFAIWIAGRPSTQRMSFGWYRAEVIGAMASVFMIWVITGILVWLAIGRLISGDYEVNA 160

Qy        216 DIMLITAAVGVAVNVIMGFLLNQ-------SGHRHSH------SHSLPSNSPTRGSGCER 262
              |||||: : : ||||||  |           || |||      ||   :::|   | :|
Db        161 KIMLITSGLAILVNVIMGVQLQHGHSHGLGGGHGHSHGGSKNASHVQATSTPCSDSPSQR 220

Qy        263 ------------------------------------------------NHGQD----- 267
                                                             :||
Db        221 IEGGVAYAPEDAELPGGGLPTFSYQNTKLVDPTLDLEIAAVLAETAPGSHHHGGPVGREA 280

Qy        268 -SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWD 326
              :: |||| :| :||::||||| :|| :| | ||| | |||||:|||::| |||| |: |
Db        281 VNMNVRAALIHVIGDVIQSVGVFVAAGVIYFWPEYSIVDPICTFVFSIIVLFTTFTIMKD 340

Qy        327 TVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSK- 385
              :::::|| |::::   : :    || |  | :| ||:|:  |      |: :   :: |
```

```
Db        341 ALLVLMEGTPNYMHYAEVLQIFQGIEGVERVHNLRIWALSINKVALSAHLAIAENANPKR 400

Qy        386 WEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANC 424
              :  : | ||  | |     |||:: |  ::: :|  |
Db        401 ILDAATSAVHLRYNFF---ETTIQIEDYTAQME-SCLQC 435


RESULT 9
Q8TC03
ID   Q8TC03      PRELIMINARY;      PRT;   388 AA.
AC   Q8TC03;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Solute carrier family 30 (Zinc transporter), member 3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC028358; AAH28358.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   388 AA;  41945 MW;  D73DE65F46860FD8 CRC64;

  Query Match           28.2%;  Score 615;  DB 4;  Length 388;
  Best Local Similarity   37.2%;  Pred. No. 2.1e-45;
  Matches  130;  Conservative   80;  Mismatches  133;  Indels    6;  Gaps    4;

Qy         79 QDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKAR--LTIAAVLYLLFMIGELVGGYIAN 136
              : ||  :  : :    |         |   :: ||  |  |  :  :|| ||:|||||:|:
Db         39 EPLPEESKPVEMPFHHCHRDPLPPPGLTPERLHARRQLYAACAVCFVFMAGEVVGGYLAH 98

Qy        137 SLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRPTFGFHRLEVLSAMISVLLVYILMG 196
              |||||||| |:| |: ::: :| :||||::  |:  |||:|| | | |: ||: :::: |
Db         99 SLAIMTDAAHLLADVGSMMGSLFSLWLSTRPATRTMTFGWHRSETLGALASVVSLWMVTG 158

Qy        197 FLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTR 256
              ||| |  | :| :| | |  ||:||:: |  |::| |:|:|:|  |||         :|
Db        159 ILLYLAFVRLLHSDYHIEGGAMLLTASIAVCANLLMAFVLHQAGPPHSHGSRGAEYAPLE 218

Qy        257 GSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLV 316
              | |      : :||||||| ||||:|| ||| |: :| |||:|| |||| |::||:
Db        219 -EGPEEPLPLGNTSVRAAFVHVLGDLLQSFGVLAASILIYFKPQYKAADPISTFLFSICA 277

Qy        317 AFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHI 376
              :|   : | : |::|| | ::  : ::: |: :  | :  :|::|:||     |  |:
Db        278 LGSTAPTLRDVLRILMEGTPRNVGFEPVRDTLLSVPGVRATHELHLWALTLTYHVASAHL 337

Qy        377 QLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQ 425
              :   |::  | | ::|:  | : ||   ||:|:: |: |: : |  ||
Db        338 AI--DSTADPEAVLAEASSRLYSRFGFSSCTLQVEQYQPEMAQ-CLRCQ 383


RESULT 10
O45923
ID   O45923      PRELIMINARY;      PRT;   410 AA.
AC   O45923;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Y39E4A.2b protein.
GN   Y39E4A.2 OR Y39E4A.2B.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Barlow K.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; AL021480; CAA16328.1; -.
DR   PIR; T26757; T26757.
DR   WormPep; Y39E4A.2b; CE16617.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   410 AA;  44082 MW;  8379EA941809270C CRC64;

  Query Match           27.8%;  Score 607.5;  DB 5;  Length 410;
  Best Local Similarity   37.7%;  Pred. No. 1e-44;
  Matches  130;  Conservative   76;  Mismatches  114;  Indels   25;  Gaps    6;

Qy         93 DSCDNC----------SKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMT 142
              ||  :|          |   :  | |: :  |   | |  :|:  | |||: | ||||||
Db         74 DSSSDCSGAGGGAHKHSHDEKYQKGRRAEKVLWAVAALSAVFIAAEFVGGFWAQSLAIMT 133

Qy        143 DALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEA 202
              || |||:|| : |::: |:  :    :|| :||: | ||| |: ||:::::|   |:  |
Db        134 DAGHMLSDLLSFIISIFAIRCARLPASKRLSFGYERAEVLGALTSVIILWVLTTVLVVVA 193

Qy        203 VQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCER 262
              :|| ::  :|:: |:||||| |||  |::|| :|:      | |:|          |
Db        194 IQRIVNNEHEVDADVMLITAGVGVLFNIVMGLVLHFGTGGHGHTH---------GGHSSH 244

Qy        263 NHGQD--SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTT 320
              |  |  :: |||| :| :|||||:|||||| ||||    : :|||||::||::| |||
Db        245 GHAHDGKNVNVRAALIHVIGDLVQSIGVLIAALIIRF-TGWTLADPICTFLFSIIVLFTT 303

Qy        321 FRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIP 380
              :: |   :::|  ||| ::  :|:||  :| |  | ||::||:   |:  ||: |
Db        304 VTVMRDIFFVLMEATPSHYDLSDVKKALSALEGVKGVHDLHLWSIGMDKTAFSVHLAL-- 361

Qy        381 GSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQ 425
              | ::  | ::|  |:   ||:   |:|:: : :::| :|  ||
Db        362 ESPNRAMENVAEARSLIRRRFGVAVATVQVEPFDEKID-SCDTCQ 405


RESULT 11
O45922
ID   O45922      PRELIMINARY;      PRT;   391 AA.
AC   O45922;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Y39E4A.2a protein.
GN   Y39E4A.2 OR Y39E4A.2A.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Barlow K.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99069613; PubMed=9851916;
```

```
RA   none;
RT   "Genome sequence of the nematode C.elegans: A platform for
RT   investigating biology.";
RL   Science 282:2012-2018(1998).
DR   EMBL; AL021480; CAA16327.1; -.
DR   PIR; T26756; T26756.
DR   WormPep; Y39E4A.2a; CE16616.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   391 AA;  42389 MW;  41F1260CEA4F6BD6 CRC64;

  Query Match           27.6%;  Score 603.5;  DB 5;  Length 391;
  Best Local Similarity  34.9%;  Pred. No. 2.2e-44;
  Matches  142;  Conservative   84;  Mismatches  140;  Indels   41;  Gaps   10;

Qy         24 NDTSAFDFSDEAGD-EGLSRFNKLRVVVADDGSEAPERPV--NGAHPTLQADDDSLLDQD 80
              ::  |  || | || ||                |     | :     |   : ||| :   :
Db         16 DENFAETFSTEEGDGEG----------GGGGDFVGRRSILTTHCHYWDENDDDMVARVE 64

Qy         81 LPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAI 140
                     :  | |: | : :    |: :  |   | |  :|:  | |||: | ||||
Db         65 ----------RGGSTDSASSRED--TGRRAEKVLWAVAALSAVFIAAEFVGGFWAQSLAI 112

Qy        141 MTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLY 200
              |||| |||:|| : |::: |:  :    :|| :||: | ||| |: ||:::::|   |:
Db        113 MTDAGHMLSDLLSFIISIFAIRCARLPASKRLSFGYERAEVLGALTSVIILWVLTTVLVV 172

Qy        201 EAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGC 260
              |:|| ::  :|:: |:|||| |||  |::|| :|:     | |:|          |
Db        173 VAIQRIVNNEHEVDADVMLITAGVGVLFNIVMGLVLHFGTGGHGHTH---------GGHS 223

Qy        261 ERNHGQD--SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAF 318
                |  |  :: |||| :| :|||||:|||||| ||||   : :||||||::||::| |
Db        224 SHGHAHDGKNVNVRAALIHVIGDLVQSIGVLIAALIIRF-TGWTLADPICTFLFSIIVLF 282

Qy        319 TTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQL 378
              || :: |   :::|  ||| ::  :|:||  :| |  | ||::||:   |:   ||: |
Db        283 TTVTVMRDIFFVLMEATPSHYDLSDVKKALSALEGVKGVHDLHLWSIGMDKTAFSVHLAL 342

Qy        379 IPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQ 425
                | ::  |  ::|  |:   ||:   |:|:: : :::| :|  ||
Db        343 --ESPNRAMENVAEARSLIRRRFGVAVATVQVEPFDEKID-SCDTCQ 386


RESULT 12
Q9VKA3
ID   Q9VKA3      PRELIMINARY;      PRT;   669 AA.
AC   Q9VKA3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   CG31860 protein.
GN   CG31860 OR CG17215.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003635; AAF53175.2; -.
DR   FlyBase; FBgn0051860; CG31860.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004182; F:carboxypeptidase A activity; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
```

```
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   InterPro; IPR000834; Peptidase_M14.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
DR   PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
SQ   SEQUENCE   669 AA;  74274 MW;  CB9C3295565FC80E CRC64;

  Query Match           27.5%;  Score 599.5;  DB 5;  Length 669;
  Best Local Similarity   29.5%;  Pred. No. 1e-43;
  Matches  143;  Conservative  103;  Mismatches  166;  Indels   73;  Gaps   13;

Qy          7 WKRLKSMLRKDDAPLFLNDTS--AFDFSDEAG---DEGLSRFNKLRVVVADDGSEAPERP 61
              |||   ::  | :     |  |  | |   :     :|      | |:|    :    ||:|
Db        159 WKRPTPVVLKSNQNYSANKNSENALDAKPQTATNTEETHGCINILKVPFQSNLYYVPEKP 218

Qy         62 VN------------GAHPTLQADDDSLLDQD------LP--LTNSQLSLKVDSCDNCS-- 99
                           || |     ::| :| |      :|  : ||:   |:|: |:
Db        219 SRQDKGSGTSDFQPGA-PETPIVENSAVDSDRKAVEIMPENVKNSE-EKKIDNSDSTKTV 276

Qy        100 -----------------KQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIM 141
                              |:||    :  :  | || :|  :|:| |::|| ::|||||
Db        277 TITGHSHITAKWDGHCHFKERETGVDKAARRVLIIACILCTIFLILEVIGGILSNSLAIA 336

Qy        142 TDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYE 201
              ||| |:||||:: :::: || |: :  ::|  :|:|| ||: ||:|:  :::: | |:|
Db        337 TDAAHLLTDLASFLISISALHLAGRPSSERLNYGWHRAEVIGAMVSIFFIWVVTGILVYM 396

Qy        202 AVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPS-NSPTRGSGC 260
              |: | :: ::|::  |||||:|: :  ||||   |     :| ||||||  :  :: :|
Db        397 AIMRWVNQDFELDAKIMLITSALAILFNVIMAMQL-----QHGHSHSLPGVHKMSKDAGS 451

Qy        261 ERN-----------------HGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKI 303
                                  | ::: |||| :| :||::|| || :|| || | ||:
Db        452 VLGSKMILLLGKSVSMQYAAKGHENINVRAAIIHVVGDIIQSFGVFVAALIIFFWPEWAF 511

Qy        304 ADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIW 363
              | :||:|||:||   ||:|: | :::::|  |  :: : :|:  :  |  |  | :| ||
Db        512 MDSVCTFVFSVLVLVVTFKILRDVLMVLMEATPDFMDYEEVKQTFLSISGVEHVHNLRIW 571

Qy        364 SLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCAN 423
              :|:  |     |: :     :  : :  :|  |:   |  :  |||:: |   ::  |
Db        572 ALSINKVALSAHLAI--SKDADPQLILEEATTLIHKRFKFFETTIQIEEYSPGME-NCGQ 628

Qy        424 CQSSS 428
              | | |
Db        629 CLSPS 633


RESULT 13
Q8T0G1
ID   Q8T0G1      PRELIMINARY;      PRT;   498 AA.
AC   Q8T0G1;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   LD05335p (CG11163-PA).
GN   BEST:CK02137 OR CG11163.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
```

```
RN   [6]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY069340; AAL39485.1; -.
DR   EMBL; AE003785; AAM68358.1; -.
DR   FlyBase; FBgn0025693; BEST:CK02137.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   498 AA;  55014 MW;  AE0161C0DC3654E8 CRC64;

  Query Match           25.9%;  Score 565.5;  DB 5;  Length 498;
  Best Local Similarity  29.3%;  Pred. No. 6.6e-41;
  Matches  134;  Conservative   91;  Mismatches  155;  Indels   77;  Gaps    8;

Qy         25 DTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPERPVNGAHPTLQADDDSLLDQDLPLT 84
Db         45 DDRLSDYQDRYYSVNNNSVNREVVTIGSNVVRVPEEQ-TWEVPLLGDCGDSARNQS---S 100

Qy         85 NSQLSLKVDSCDNCSKQREILKQR---KVKARLTIAAVLYLLFMIGELVGGYIANSLAIM 141
Db        101 ETEFDMQRECCNHQPGFRANSKSKSAQEAKYKIMLAVALCCVFMIIEFLGGYVAGSLAIM 160

Qy        142 TDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYE 201
Db        161 TDAAHLASDCISFVIGLVAIWIGGRPPDERMSFGYKRFEVIGALASILGIWFVTTLLVVV 220

Qy        202 AVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLL------NQSGHRHSHSHSLPSNSPT 255
Db        221 AIERIFSQDFELNADMMMLISGIGIVINIVMMFVLHGSWFVNGNGHGHSHSHS-HSHSHS 279

Qy        256 RGSGCERN---------------------------------------------------- 263
Db        280 HGNGHEPNDSLSQTRSNSNFLTTIGSQSASTADEEDSIRKEINSNEHKIVITNGKKPTLT 339

Qy        264 --------HGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLL 315
Db        340 GTSNMELAHEDKNLNLRAAMIHVIGDLVQSIGVFLAAVLIKVCPGAKYADPLCTLIFSII 399

Qy        316 VAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVH 375
Db        400 VIMTTLRLFRESLGIIVNAVPQNLNMRTLHLELGSIEGVRSLHHLNVWQQTSQQRVLMVH 459

Qy        376 IQLIPGSSSKWEEVQSKANHLLLN-TFGMYRCTIQLQ 411
Db        460 --LVTDSRADGNEVLQAATALVSSPRYNIKHSTIQLE 494


RESULT 14
O81036
ID   O81036      PRELIMINARY;      PRT;   398 AA.
AC   O81036;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Putative zinc transporter (ZAT).
GN   AT2G46800.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=20083487; PubMed=10617197;
RA   Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA   Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA   Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA   Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA   Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA   Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA   Salzberg S.L., Fraser C.M., Venter J.C.;
RT   "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:761-768(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Lin X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT   "Full-length messenger RNA sequences greatly improve genome
RT   annotation.";
RL   Genome Biol. 0:0-0(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RT   "Full-Length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC005310; AAC33498.1; -.
DR   EMBL; AY086033; AAM63243.1; -.
DR   PIR; T02681; T02681.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   398 AA;  43827 MW;  7E20E0B29237BB23 CRC64;

  Query Match           25.4%;  Score 555;  DB 10;  Length 398;
  Best Local Similarity  33.8%;  Pred. No. 4.1e-40;
  Matches  116;  Conservative   74;  Mismatches  107;  Indels   46;  Gaps    4;

Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRF 172
Db         57 KLCIAVVLCLVFMSVEVVGGIKANSLAILTDAAHLLSDVAAFAISLFSLWAAGWEATPRQ 116

Qy        173 TFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIM 232
Db        117 TYGFFRIEILGALVSIQLIWLLTGILVYEAIIRIVTETSEVNGFLMFLVAAFGLVVNIIM 176

Qy        233 GFLLNQS-------GHRHSHSHSLPSNSPT---------RGSGCERNHGQD--------- 267
Db        177 AVLLGHDHGHSHGHGHGHGHDHHNHSHGVTVTTHHHHHDHEHGHSHGHGEDKHHAHGDVT 236

Qy        268 ------------------SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPIC 308
Db        237 EQLLDKSKTQVAAKEKRKRNINLQGAYLHVLGDSIQSVGVMIGGAIIWYNPEWKIVDLIC 296

Qy        309 TYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSG 368
Db        297 TLAFSVIVLGTTINMIRNILEVLMESTPREIDATKLEKGLLEMEEVVAVHELHIWAITVG 356

Qy        369 KSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
Db        357 KVLLACHVNIRPEADA--DMVLNKVIDYIRREYNISHVTIQIE 397


RESULT 15
Q9ZT63
ID   Q9ZT63      PRELIMINARY;      PRT;   398 AA.
AC   Q9ZT63;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Zinc transporter ZAT.
```

```
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. ecotype C24;
RX   MEDLINE=99169295; PubMed=10069843;
RA   van der Zaal B.J., Neuteboom L.W., Pinas J.E., Chardonnens A.N.,
RA   Schat H., Verkleij J.A.C., Hooykaas P.J.J.;
RT   "Overexpression of a novel arabidopsis gene related to putative zinc-
RT   transporter genes from animals can lead to enhanced zinc resistance
RT   and accumulation.";
RL   Plant Physiol. 119:1047-1055(1999).
DR   EMBL; AF072858; AAD11757.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0008324; F:cation transporter activity; IEA.
DR   GO; GO:0006812; P:cation transport; IEA.
DR   InterPro; IPR002524; Cation_efflux.
DR   Pfam; PF01545; Cation_efflux; 1.
DR   TIGRFAMs; TIGR01297; CDF; 1.
SQ   SEQUENCE   398 AA;  44063 MW;  6FC33112236D7C3C CRC64;

  Query Match             25.1%;  Score 547;  DB 10;  Length 398;
  Best Local Similarity   33.2%;  Pred. No. 2e-39;
  Matches  114;  Conservative   80;  Mismatches  103;  Indels   46;  Gaps    5;

Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRF 172
Db         57 KLCIAVVLCLVFMSVEVVGGIKANSLAILTDAAHLLSDVAAFAISLFSLWAAGWEATPRQ 116

Qy        173 TFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIM 232
Db        117 TYGFFRIEILGALVSIQLIWLLTGILVYEAIIRIVTETSEVNGFLMFLVAAFGLVVNIIM 176

Qy        233 GFLLNQS-----------GH-RHSHSHSLP----------SNSPTRGSGCERNHG----- 265
Db        177 AVLLGHDHGPQSWTWAWHGHDHHNHSHGVTVTTHHHHHDHEHGHSHGHGEDKHHAHGDVT 236

Qy        266 ----------------QDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPIC 308
Db        237 EQLLDKSKTQVAAKEKRKRNINLQGAYLHVLGDSIQSVGVMIGGAIIWYNPEWKIVDLIC 296

Qy        309 TYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSG 368
Db        297 TLAFSVIVLGTTINMIRNILEVLMESTPREIDATKLEKGLLEMEEVVAVHELHIWAITVG 356

Qy        369 KSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
Db        357 KVLLACHVNIRPEADA--DMVLNKVIDYIRREYNISHVTIQIE 397


Search completed: July 29, 2004, 17:44:24
Job time : 42 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 29, 2004, 17:41:15 ; Search time 17 Seconds
(without alignments)
2427.421 Million cell updates/sec

Title: US-10-024-652-2570
Perfect score: 2183
Sequence: 1 MAGSGAWKRLKSMLRKDDAP.....LQSYRQEVDRTCANCQSSSP 429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	747	34.2	359	1	S70632	zinc transporter Z
2	638.5	29.2	382	2	T24963	hypothetical prote
3	607.5	27.8	410	2	T26757	hypothetical prote
4	603.5	27.6	391	2	T26756	hypothetical prote
5	555	25.4	398	2	T02681	probable zinc tran
6	534	24.5	385	2	A84696	probable zinc tran
7	532	24.4	378	2	T49164	zinc transporter-l
8	526.5	24.1	334	2	T47986	zinc transporter-l
9	437.5	20.0	319	2	D89778	hypothetical prote
10	407	18.6	308	2	E70392	cation efflux syst
11	391	17.9	303	2	AG1396	cation transport p
12	387	17.7	303	2	AB1772	cation transport p
13	386	17.7	313	1	H64810	ybgR protein - Esc
14	385	17.6	311	2	E85577	probable transport
15	385	17.6	313	2	D90726	probable transport
16	372	17.0	312	2	AE0593	probable cation tr
17	370	16.9	311	2	C69612	cation-efflux syst
18	369.5	16.9	325	2	T44365	cation-efflux syst
19	369.5	16.9	325	2	G90008	hypothetical prote
20	368	16.9	299	2	C83595	probable cation ef
21	366.5	16.8	312	2	T35276	probable efflux pr
22	357.5	16.4	301	2	AH3431	cobalt-zinc-cadmiu
23	355	16.3	312	2	AI0138	probable cation tr
24	353	16.2	316	1	JC4701	cadmium, zinc, cob
25	351.5	16.1	321	2	G82752	cobalt-zinc-cadmiu
26	343	15.7	387	2	T38252	probable zinc/cadm
27	343	15.7	387	2	T43140	probable heavy met
28	343	15.7	439	2	S58327	cobalt accumulatio
29	342	15.7	442	2	S56057	heavy metal ion re
30	341	15.6	303	2	H97212	co/Zn/Cd efflux sy
31	337	15.4	325	2	A75437	cation efflux syst
32	336.5	15.4	300	2	D84459	probable cation tr
33	325.5	14.9	284	2	G72363	cation efflux syst
34	314.5	14.4	304	2	AG2540	cation efflux syst
35	306.5	14.0	316	2	E81321	probable cation tr
36	296	13.6	295	2	H71078	probable cation ef
37	284.5	13.0	507	2	S54303	zinc transport pro
38	282.5	12.9	299	2	F98080	hypothetical prote
39	282.5	12.9	299	2	H95216	cation efflux syst
40	282.5	12.9	503	2	S54302	zinc transporter Z
41	271.5	12.4	361	2	F87286	cation efflux fami
42	264.5	12.1	304	2	G86825	cation transport p
43	249	11.4	199	1	D33830	cation efflux syst
44	248	11.4	281	2	AF2161	cation-efflux syst
45	245.5	11.2	326	2	D83483	probable metal tra

ALIGNMENTS

RESULT 1
S70632
zinc transporter ZnT-2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: S70632
R;Palmiter, R.D.; Cole, T.B.; Findley, S.D.
EMBO J. 15, 1784-1791, 1996
A;Title: ZnT-2, a mammalian protein that confers resistance to zinc by facilitating ves
A;Reference number: S70632; MUID:96203098; PMID:8617223
A;Accession: S70632
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-359 <PAL>
A;Cross-references: EMBL:U50927; NID:g1256377; PIDN:AAB02775.1; PID:g1256378
C;Genetics:
A;Gene: ZnT-2
A;Start codon: CTG
C;Superfamily: zinc transporter ZnT-2

Query Match 34.2%; Score 747; DB 1; Length 359;
Best Local Similarity 42.0%; Pred. No. 5.9e-53;
Matches 148; Conservative 78; Mismatches 108; Indels 18; Gaps 4;

```
Qy         82 PLTNSQLSLKVDSCDNCSKQREI-----LKQRKVKARLTIAAVLYLLFMIGELVGGYIAN 136
              |:    : | | |    |  |::       :::: : :| :|: : |:|||||::|||:|
Db         20 PVNLPSVELAVQSNHYCHAQKDSGSHPNSEKQRARRKLYVASAICLVFMIGEIIGGYLAQ 79

Qy        137 SLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMG 196
              ||||||| |:||| :::::::| :||:||:  ||   ||: | |:| |::||| :::: |
Db         80 SLAIMTDAAHLLTDFASMLISLFSLWVSSRPATKTMNFGWQRAEILGALLSVLSIWVVTG 139

Qy        197 FLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTR 256
              |:| |||| |  :||| || ||||:   ||||:|||  |:|||| ||| ||   :|
Db        140 VLVYLAVQRLISGDYEIKGDTMLITSGCAVAVNIIMGLALHQSGHGHSHGHSHEDSS--- 196

Qy        257 GSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLV 316
                    | : :|||||:| :|||:||||||:||||| ||||||  |||||::||:||
Db        197 -------QQQQNPSVRAAFIHVVGDLLQSVGVLVAAYIIYFKPEYKYVDPICTFLFSILV 249

Qy        317 AFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHI 376
               ||  |: | :::::|| |  ::   :|  |: :: | ::  |:||:|| :    |||
Db        250 LGTTLTILRDVILVLMEGTPKGVDFTTVKNLLLSVDGVEALHSLHIWALTVAQPVLSVHI 309

Qy        377 QLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQSSS 428
              :    :  : |   |   |   |  :  |||::|| ::: ::|  ||  |
Db        310 AIAQNVDA--QAVLKVARDRLQGKFNFHTMTIQIESYSEDM-KSCQECQGPS 358
```

RESULT 2
T24963
hypothetical protein T18D3.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C;Accession: T24963
R;McMurray, A.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19962
A;Accession: T24963
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-382 <WIL>
A;Cross-references: EMBL:Z68119; PIDN:CAA92193.1; GSPDB:GN00028; CESP:T18D3.3
A;Experimental source: clone T18D3
C;Genetics:
A;Gene: CESP:T18D3.3
A;Map position: X
A;Introns: 34/2; 72/1; 142/1; 193/2; 254/3; 290/1; 325/3

Query Match 29.2%; Score 638.5; DB 2; Length 382;
Best Local Similarity 39.8%; Pred. No. 4.1e-44;
Matches 136; Conservative 70; Mismatches 105; Indels 31; Gaps 8;

```
Qy          91 KVDSCDNCSKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTD 150
               :  || |:           |:     | :  || | ||: |::|| :| ||||:||| |:|||
Db          59 EADSTDSHD------SNRRATRILWLTVVLCLFFMVCEVIGGVLAGSLAIVTDAAHLLTD 112

Qy         151 LSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMN 210
                ::::::| :|::: : |::: :||||| ||| |  || |::|: | |:  |: | :  :
Db         113 FASVLISLFSLYIARRPPSQKMSFGFHRAEVLGAFFSVFLIWIVTGVLVVLAIMRIVSGD 172

Qy         211 YEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQ---- 266
               ||: | || :|||:|| ||::|  ||    || |||              | |    |
Db         173 YEVEGGIMALTAALGVVVNLVMLALLYFGGHSHSH----------GGGSSHGHSHGGGN 221

Qy         267 -DSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIW 325
               |:: |||||:| ||||:||:|||:||  | |:| : | ||||| |||::|  ||  |:
Db         222 GDNINVRAAFIHVLGDLLQSLGVLVAALFIYFQPSWVIIDPICTLVFSVIVLCTTIYILR 281

Qy         326 DTVVIILEGVPSHLNVDYIK--EALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSS 383
               | ::::||| ||  |:|: |   :|  || |  | || |||||  |    ||:::   |
Db         282 DAMIVLLEGRPS--NIDFAKVFSSLEDIEGVKKVHDLRIWSLTMDKIALSVHLEIDANSQ 339

Qy         384 SKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT-CANC 424
               |  : :  :   :|  |: ::  |||:: :    :|: | |
Db         340 S--QSILRETRKMLKQTYNVHEITIQIEEF--GANRSDCGKC 377
```

RESULT 3
T26757
hypothetical protein Y39E4A.2b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26757
R;Barlow, K.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z20260
A;Accession: T26757
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-410 <WIL>
A;Cross-references: EMBL:AL021480; PIDN:CAA16328.1; GSPDB:GN00021; CESP:Y39E4A.2b
A;Experimental source: clone Y39E4A
C;Genetics:
A;Gene: CESP:Y39E4A.2b
A;Map position: 3
A;Introns: 97/3; 362/1

Query Match 27.8%; Score 607.5; DB 2; Length 410;
Best Local Similarity 37.7%; Pred. No. 1.5e-41;
Matches 130; Conservative 76; Mismatches 114; Indels 25; Gaps 6;

```
Qy          93 DSCDNC----------SKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMT 142
               ||  :|          |    :  | |: :  |    | |  :|:  | |||: | ||||||
Db          74 DSSSDCSGAGGGAHKHSHDEKYQKGRRAEKVLWAVAALSAVFIAAEFVGGFWAQSLAIMT 133

Qy         143 DALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEA 202
               || |||:|| : |::: |:  :    :|| :||: | |||  |: ||:::::|   |:  |
Db         134 DAGHMLSDLLSFIISIFAIRCARLPASKRLSFGYERAEVLGALTSVIILWVLTTVLVVVA 193

Qy         203 VQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCER 262
               :|| ::  :|:: |:|||| |||  |::|| :|:     | |:|          |
Db         194 IQRIVNNEHEVDADVMLITAGVGVLFNIVMGLVLHFGTGGHGHTH---------GGHSSH 244

Qy         263 NHGQD--SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTT 320
                 |  |  :: |||| :| :|||||:|||||| |||   : :|||||::||::| |||
Db         245 GHAHDGKNVNVRAALIHVIGDLVQSIGVLIAALIIRF-TGWTLADPICTFLFSIIVLFTT 303

Qy         321 FRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIP 380
               :: |   :::|  ||| ::  :|:||  :| |  | ||::||:   |:   ||: |
Db         304 VTVMRDIFFVLMEATPSHYDLSDVKKALSALEGVKGVHDLHLWSIGMDKTAFSVHLAL-- 361

Qy         381 GSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQ 425
                | ::  |  ::|  |:   ||:   |:|:: : :::| :|  ||
Db         362 ESPNRAMENVAEARSLIRRRFGVAVATVQVEPFDEKID-SCDTCQ 405
```

RESULT 4
T26756
hypothetical protein Y39E4A.2a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T26756
R;Barlow, K.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z20260
A;Accession: T26756
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-391 <WIL>
A;Cross-references: EMBL:AL021480; PIDN:CAA16327.1; GSPDB:GN00021; CESP:Y39E4A.2a
A;Experimental source: clone Y39E4A
C;Genetics:
A;Gene: CESP:Y39E4A.2a
A;Map position: 3
A;Introns: 22/1; 78/3; 343/1

Query Match 27.6%; Score 603.5; DB 2; Length 391;
Best Local Similarity 34.9%; Pred. No. 2.9e-41;
Matches 142; Conservative 84; Mismatches 140; Indels 41; Gaps 10;

```
Qy          24 NDTSAFDFSDEAGD-EGLSRFNKLRVVVADDGSEAPERPV--NGAHPTLQADDDSLLDQD 80
               ::  |  || | || ||                |     | :     |   : ||| :   :
Db          16 DENFAETFSTEEGDGEG----------GGGGDFVGRRSILTTHCHYWDENDDDMVARVE 64

Qy          81 LPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAI 140
                         :  | |: | : :    |: :  |   | |  :|:  | ||||: | ||||
Db          65 ----------RGGSTDSASSRED--TGRRAEKVLWAVAALSAVFIAAEFVGGFWAQSLAI 112

Qy         141 MTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLY 200
               |||| |||:|| : |::: |:  :    :|| :||: | ||| |: ||:::::|   |:
Db         113 MTDAGHMLSDLLSFIISIFAIRCARLPASKRLSFGYERAEVLGALTSVIILWVLTTVLVV 172

Qy         201 EAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGC 260
               |:|| ::  :|:: |:|||| |||  |::|| :|:     | |:|          |
Db         173 VAIQRIVNNEHEVDADVMLITAGVGVLFNIVMGLVLHFGTGGHGHTH---------GGHS 223

Qy         261 ERNHGQD--SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAF 318
                  |  |  :: |||| :| :||||||:|||||| ||||   : :|||||::||::| |
Db         224 SHGHAHDGKNVNVRAALIHVIGDLVQSIGVLIAALIIRF-TGWTLADPICTFLFSIIVLF 282

Qy         319 TTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQL 378
               ||  :: |   :::|  ||| ::  :|:||  :| |  | ||::||:   |:   ||: |
Db         283 TTVTVMRDIFFVLMEATPSHYDLSDVKKALSALEGVKGVHDLHLWSIGMDKTAFSVHLAL 342

Qy         379 IPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQ 425
                | ::  |  ::|  |:   ||:   |:|:: : :::| :|  ||
Db         343 --ESPNRAMENVAEARSLIRRRFGVAVATVQVEPFDEKID-SCDTCQ 386
```

RESULT 5
T02681
probable zinc transporter At2g46800 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F19D11.8
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Feb-2001
C;Accession: T02681; D84907
R;Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Ron
submitted to the EMBL Data Library, September 1998
A;Description: Arabidopsis thaliana chromosome II BAC F19D11 genomic sequence.
A;Reference number: Z14698
A;Accession: T02681
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-398 <ROU>
A;Cross-references: EMBL:AC005310; NID:g3510247; PID:g3510254
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: D84907
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-398 <STO>
A;Cross-references: GB:AE002093; NID:g3510254; PIDN:AAC33498.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g46800; F19D11.8
A;Map position: 2

  Query Match 25.4%; Score 555; DB 2; Length 398;
  Best Local Similarity 33.8%; Pred. No. 2.6e-37;
  Matches 116; Conservative 74; Mismatches 107; Indels 46; Gaps 4;

```
Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRF 172
Db         57 KLCIAVVLCLVFMSVEVVGGIKANSLAILTDAAHLLSDVAAFAISLFSLWAAGWEATPRQ 116

Qy        173 TFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIM 232
Db        117 TYGFFRIEILGALVSIQLIWLLTGILVYEAIIRIVTETSEVNGFLMFLVAAFGLVVNIIM 176

Qy        233 GFLLNQS-------GHRHSHSHSLPSNSPT---------RGSGCERNHGQD--------- 267
Db        177 AVLLGHDHGHSHGHGHGHGHDHHNHSHGVTVTTHHHHHDHEHGHSHGHGEDKHHAHGDVT 236

Qy        268 ------------------SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPIC 308
Db        237 EQLLDKSKTQVAAKEKRKRNINLQGAYLHVLGDSIQSVGVMIGGAIIWYNPEWKIVDLIC 296

Qy        309 TYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSG 368
Db        297 TLAFSVIVLGTTINMIRNILEVLMESTPREIDATKLEKGLLEMEEVVAVHELHIWAITVG 356

Qy        369 KSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
Db        357 KVLLACHVNIRPEADA--DMVLNKVIDYIRREYNISHVTIQIE 397
```

RESULT 6
A84696
probable zinc transporter [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: A84696
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: A84696
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-385 <STO>
A;Cross-references: GB:AE002093; NID:g3980394; PIDN:AAC95197.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g29410
A;Map position: 2

  Query Match 24.5%; Score 534; DB 2; Length 385;
  Best Local Similarity 32.4%; Pred. No. 1.3e-35;
  Matches 122; Conservative 92; Mismatches 120; Indels 42; Gaps 9;

```
Qy         69 LQADDDSLLDQDLP-LTNSQLSLKVDSC-----DNC---SKQREILKQRKVKARLTIAAV 119
Db         19 LKPDDEEEMESPSPSKTEENLGVVPLSCAFTRQEHCVSETKERE-----ESTRRLSSLIF 73

Qy        120 LYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRL 179
Db         74 LYLIVMSVQIVGGFKANSLAVMTDAAHLLSDVAGLCVSLLAIKVSSWEANPRNSFGFKRL 133

Qy        180 EVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLL--N 237
Db        134 EVLAAFLSVQLIWLVSGVIIHEAIQRLLSRSREVNGEIMFGISAFGFFMNLVMVLWLGHN 193

Qy        238 QSGHRHSHSHSLPSN-------------------SPTRGSGCERNHGQD-SLAVRAAFV 276
Db        194 HSHHHHDHHHHHHNHKHQHQHHHKEVVAEEEEEEMNPLKG---EKSSSKEMNINIQGAYL 250

Qy        277 HALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVP 336
Db        251 HAMADMIQSLGVMIGGGIIWVKPKWVLVDLICTLVFSAFALAATLPILKNIFGILMERVP 310

Qy        337 SHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHL 396
Db        311 RDMDIEKLERGLKRIDGVKIVYDLHVWEITVGRIVLSCHILPEPGASPK--EIITGVRNF 368

Qy        397 LLNTFGMYRCTIQLQS 412
Db        369 CRKSYGIYHATVQVES 384
```

RESULT 7
T49164
zinc transporter-like protein - Arabidopsis thaliana
N;Alternate names: protein T20N10.160
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49164
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Rudd, S.; I
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25017
A;Accession: T49164
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-378 <DAN>
A;Cross-references: EMBL:AL353032; GSPDB:GN00061; ATSP:T20N10.160
A;Experimental source: cultivar Columbia; BAC clone T20N10
C;Genetics:
A;Gene: ATSP:T20N10.160
A;Map position: 3

  Query Match 24.4%; Score 532; DB 2; Length 378;
  Best Local Similarity 33.6%; Pred. No. 1.8e-35;
  Matches 108; Conservative 79; Mismatches 110; Indels 24; Gaps 3;

```
Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRF 172
Db         58 KLLIAVLLCAIFIVVEVVGGIKANSLAILTDAAHLLSDVAAFAISLFSLWASGWKANPQQ 117

Qy        173 TFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIM 232
```

```
Db        118 SYGFFRIEILGALVSIQMIWLLAGILVYEAIVRLNNGSGEVEGSLMFAVSAVGLLVNIAM 177

Qy        233 GFLL-------------NQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLA--------- 270

Db        178 AILLGHDHGHGHGHSHDNGHGHSHDHGHGIAATEHHHDSGHDESQLSDVLIEQKKQRNVN 237

Qy        271 VRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVI 330

Db        238 IQGAYLHVLGDSIQSVGVMIGGAIIWYKPEWKILDLICTLVFSVIVLGTTIGMLRNILEV 297

Qy        331 ILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQ 390

Db        298 LMESTPREIDPTMLEKGVCEIEEVVAVHELHIWAITVGKLLLACHVKIRPEAEA--DMVL 355

Qy        391 SKANHLLLNTFGMYRCTIQLQ 411

Db        356 DKIIDYIKREHNISHVTIQIE 376


RESULT 8
T47986
zinc transporter-like protein - Arabidopsis thaliana
N;Alternate names: protein F21F14.110
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 27-Nov-2001
C;Accession: T47986
R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24481
A;Accession: T47986
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-334 <CHO>
A;Cross-references: EMBL:AL138642
A;Experimental source: cultivar Columbia; BAC clone F21F14
C;Genetics:
A;Map position: 3
A;Note: F21F14.110
C;Superfamily: zinc transporter ZnT-2

  Query Match             24.1%;  Score 526.5;  DB 2;  Length 334;
  Best Local Similarity   34.4%;  Pred. No. 4.3e-35;
  Matches  111;  Conservative   80;  Mismatches  125;  Indels    7;  Gaps    4;

Qy         94 SCDNCSKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSA 153

Db         13 ACGFSTSSSDAKKRAASMRKLCFVVVLCLLFMSIEVVCGIKANSLAILADAAHLLTDVGA 72

Qy        154 IILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTI-HMNYE 212

Db         73 FAISMLSLWASSWEANPRQSYGFFRIEILGTLVSIQLIWLLTGILVYEAVTRLVQETNDD 132

Qy        213 INGDIMLITAAVGVAVNVIMGFLLNQS---GHRHSHSHSLPSNSPTRGSG-CERNHGQDS 268

Db        133 VDGFFMVLVAAFGLVVNIIMIVVLGHDHGHGHDHGHSHDHGHSYGERAEQLLEKSKEIRN 192

Qy        269 LAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTV 328

Db        193 INVQGAYLHVLGDLIQSIGVMIGGGMIWYNPKWKVIDLICTLFFSVIVLGTTIKMLRSIL 252

Qy        329 VIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEE 388

Db        253 EVLMESTPREIDARQLEKGLMEIEEVVDVHELHIWAITVGKALFSCHVKVRPEAGD--EM 310

Qy        389 VQSKANHLLLNTFGMYRCTIQLQ 411

Db        311 VLNKVIDYIWREYRISHVTIQIE 333


RESULT 9
D89778
hypothetical protein SA0163 [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: D89778
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: D89778
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <KUR>
A;Cross-references: GB:BA000018; PID:g13700084; PIDN:BAB41383.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: SA0163
C;Superfamily: zinc transporter ZnT-2

  Query Match             20.0%;  Score 437.5;  DB 2;  Length 319;
  Best Local Similarity   30.3%;  Pred. No. 6.9e-28;
  Matches   94;  Conservative   84;  Mismatches   99;  Indels   33;  Gaps    5;

Qy        107 QRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSK 166

Db         24 QQSSKKTLWASLIITLLFTVIEFVGGLVSNSLALLSDSFHMLSDVLALGLSMLAIYFASK 83

Qy        167 SPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGV 226

Db         84 KPTARYTFGYLRFEILVAFLNGLALIVISIWILYEAIVRIIYPQ-PIESGIMFMIASIGL 142

Qy        227 AVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSV 286

Db        143 LVNIILTIILVRSLKQ----------------------EDNINIQSALWHFMGDLLNSI 179

Qy        287 GVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKE 346

Db        180 GVIVAVVLIYF-TGWRIIDPIISIVISLIILRGGYKITRNAWLILMESVPQHLDTDQIMA 238

Qy        347 ALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWE----EVQSKANHLLLNTFG 402

Db        239 DIKNIDGILDVHEFHLWSITTEHYSLSAHVVL----DKKYEGDDYQAIDQVSSLLKEKYG 294

Qy        403 MYRCTIQLQS 412

Db        295 IAHSTLQIEN 304


RESULT 10
E70392
cation efflux system (czcD-like) - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 20-Sep-1999
C;Accession: E70392
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: E70392
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-308 <AQF>
A;Cross-references: GB:AE000721; NID:g2983544; PIDN:AAC07126.1; PID:g2983552; GB:AE00065
A;Experimental source: strain VF5
C;Genetics:
A;Gene: czcD
C;Superfamily: zinc transporter ZnT-2

  Query Match             18.6%;  Score 407;  DB 2;  Length 308;
  Best Local Similarity   29.5%;  Pred. No. 2e-25;
  Matches   96;  Conservative   77;  Mismatches  104;  Indels   48;  Gaps    7;
```

```
Qy        106 KQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSS 165
Db          3 REKSLKV-LAFSFLLIFLFAFIEFLGGLLTNSLALLSDAGHMLTDAVSLSIALVAQYLAL 61

Qy        166 KSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVG 225
Db         62 KVKTKRTTYGLYRLEVLAALVNGVFLLGLIGYIILEAIHR--FENPEPVKPQMIYIAFAG 119

Qy        226 VAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQS 285
Db        120 LIVNLVVGYIL-----------------------LKHSEENINIKSALLHVATDTLGS 154

Qy        286 VGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIK 345
Db        155 VAAIIAGIAIVFWKFY-LADPILSVAVALLILPSAYSVIKETVNVLLEVAPSHINTEELE 213

Qy        346 EALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLL------LN 399
Db        214 KELLNLQGVKGVHDLHVWSITPGTEVLTVHVVV---------EDTSICNDILKEVEKIAH 264

Qy        400 TFGMYRCTIQLQSYRQEVDRTCANC 424
Db        265 KYGIKHTTVQL----EKEGYACAEC 285


RESULT 11
AG1396
cation transport protein (efflux) homolog lmo2575 [imported] - Listeria monocytogenes (s
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AG1396
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AG1396
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-303 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00653.1; PID:g16412063; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2575
C;Superfamily: zinc transporter ZnT-2

  Query Match             17.9%;  Score 391;  DB 2;  Length 303;
  Best Local Similarity   29.2%;  Pred. No. 3.8e-24;
  Matches   88;  Conservative   76;  Mismatches  109;  Indels   28;  Gaps    4;

Qy        111 KARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTK 170
Db         21 KKSLFISFILIATFMVVEVIGGIMTNSLALLSDAGHMLSDAVALGLSLAAFKFGEKAASS 80

Qy        171 RFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNV 230
Db         81 DKTYGYKRFEILAAFLNGLTLVGISVFIFYEAIGRFFDPPQVIGAGMMTI-SVIGLLINI 139

Qy        231 IMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLI 290
Db        140 LVAWIL-----------------------MKGDTSENLNMRSAFLHVLGDLLGSVGAIT 175

Qy        291 AAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK 350
Db        176 AALLIIFL-GWNIADPIASVIVAALILVSGWRVLKDAIHILMEGKPANVDTEEIKTFFQQ 234

Qy        351 IEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQL 410
Db        235 QDGVKEVHDLHVWAITSDFNALSAHLTVCEDADR--DKILADIEHYLQENFSLEHSTIQL 292


RESULT 12
AB1772
cation transport protein (efflux) homolog lin2720 [imported] - Listeria innocua (strain
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C;Accession: AB1772
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1772
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-303 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97946.1; PID:g16415256; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2720
C;Superfamily: zinc transporter ZnT-2

  Query Match             17.7%;  Score 387;  DB 2;  Length 303;
  Best Local Similarity   29.2%;  Pred. No. 8.1e-24;
  Matches   88;  Conservative   75;  Mismatches  110;  Indels   28;  Gaps    4;

Qy        111 KARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTK 170
Db         21 KKSLFISFILIATFMIVEVIGGIMTNSLALLSDAGHMLSDAVALGLSLAAFKFGEKAASS 80

Qy        171 RFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNV 230
Db         81 DKTYGYKRFEILAAFLNGLTLVGISVFIFYEAIGRFFDPPQVIGAGMMTI-SVIGLLINI 139

Qy        231 IMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLI 290
Db        140 LVAWIL-----------------------MKGDTSENLNMRSAFLHVLGDLLGSVGAIT 175

Qy        291 AAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK 350
Db        176 AALLIIFF-GWNIADPIASVIVAALILVSGWRVLKDAIHILMEGKPANVDTEEIKTFFQE 234

Qy        351 IEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQL 410
Db        235 QDGVTEVHDLHVWAITSDFNALTAHLTV--AEDADRDKILTDIEHYLQENFSLEHSTIQL 292

Qy        411 Q 411
Db        293 E 293


RESULT 13
H64810
ybgR protein - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Mar-2002
C;Accession: H64810
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: H64810
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-313 <BLAT>
```

```
A;Cross-references: GB:AE000177; GB:U00096; NID:g1786955; PIDN:AAC73839.1; PID:g1786966;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ybgR
C;Superfamily: zinc transporter ZnT-2
C;Keywords: transmembrane protein
F;21-37/Domain: transmembrane #status predicted <TM1>
F;54-70/Domain: transmembrane #status predicted <TM2>
F;90-106/Domain: transmembrane #status predicted <TM3>
F;126-142/Domain: transmembrane #status predicted <TM4>
F;164-180/Domain: transmembrane #status predicted <TM5>
F;188-204/Domain: transmembrane #status predicted <TM6>

  Query Match             17.7%;  Score 386;  DB 1;  Length 313;
  Best Local Similarity   31.6%;  Pred. No. 1e-23;
  Matches   92;  Conservative   71;  Mismatches   94;  Indels   34;  Gaps    7;

Qy        124 FMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLS 183
              ||: |:|||::: |||:: || ||||| :|:: |||:  | : || | |||: ||  |:
Db         31 FMLVEVVGGFLSGSLALLADAGHMLTDTAALLFALLAVQFSRRPPTIRHTFGWLRLTTLA 90

Qy        184 AMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRH 243
              | :: : : ::   :::||::|       : | :|:  |  |:  |::  :||
Db         91 AFVNAIALVVITILIVWEAIER-FRTPRPVEGGMMMAIAVAGLLANILSFWLL------- 142

Qy        244 SHSHSLPSNSPTRGSGCERNHGQD--SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEY 301
                              :|| :  :| |||| :| ||||: ||| :||| || :   :
Db        143 ------------------HHGSEEKNLNVRAAALHVLGDLLGSVGAIIAALIIIW-TGW 182

Qy        302 KIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK-IEDVYSVEDL 360
              ||||  : : |||  : :|:: |:|  :||| |  |::  :|  : : | :| :|  :
Db        183 TPADPILSILVSLLVLRSAWRLLKDSVNELLEGAPVSLDIAELKRRMCREIPEVRNVHHV 242

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              ::|  :   |    :|:|:||      : :  :  | |:: : :   |||::
Db        243 HVW-MVGEKPVMTLHVQVIPPHDH--DALLDQIQHYLMDHYQIEHATIQME 290


RESULT 14
E85577
probable transport system permease protein ybgR [imported] - Escherichia coli (strain O1
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: E85577
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: E85577
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <STO>
A;Cross-references: GB:AE005174; NID:g12513681; PIDN:AAG55081.1; GSPDB:GN00145; UWGP:Z09
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ybgR
C;Superfamily: zinc transporter ZnT-2

  Query Match             17.6%;  Score 385;  DB 2;  Length 311;
  Best Local Similarity   31.3%;  Pred. No. 1.2e-23;
  Matches   91;  Conservative   72;  Mismatches   94;  Indels   34;  Gaps    7;

Qy        124 FMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLS 183
              ||: |::||::: |||:: || ||||| :|::  |||:  | : || | |||: ||  |:
Db         29 FMLVEVIGGFLSGSLALLADAGHMLTDTAALLFALLAVQFSRRPPTIRHTFGWLRLTTLA 88

Qy        184 AMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRH 243
              | :: : : ::   :::||::|       : | :|:  |  |:  |::  :||
Db         89 AFVNAIALVVITILIVWEAIER-FRTPRPVEGGMMMAIAVAGLLANILSFWLL------- 140

Qy        244 SHSHSLPSNSPTRGSGCERNHGQD--SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEY 301
                              :|| :  :| |||| :| ||||: ||| :||| || :   :
Db        141 ------------------HHGSEEKNLNVRAAALHVLGDLLGSVGAIIAALIIIW-TGW 180

Qy        302 KIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK-IEDVYSVEDL 360
              ||||  : : |||  : :|:: |:|  :||| |  |::  :|  : : | :| :|  :
Db        181 TPADPILSILVSLLVLRSAWRLLKDSVNELLEGAPVSLDIAELKRRMCREIPEVRNVHHV 240

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              ::|  :   |    :|:|:||      : :  :  | |:: : :   |||::
Db        241 HVW-MVGEKPVMTLHVQVIPPHDH--DALLDQIQHYLMDHYQIEHATIQME 288


RESULT 15
D90726
probable transport system permease protein ECs0780 [imported] - Escherichia coli (strain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: D90726
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D90726
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-313 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB34203.1; PID:g13360239; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs0780
C;Superfamily: zinc transporter ZnT-2

  Query Match             17.6%;  Score 385;  DB 2;  Length 313;
  Best Local Similarity   31.3%;  Pred. No. 1.2e-23;
  Matches   91;  Conservative   72;  Mismatches   94;  Indels   34;  Gaps    7;

Qy        124 FMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLS 183
              ||: |::||::: |||:: || ||||| :|::  |||:  | : || | |||: ||  |:
Db         31 FMLVEVIGGFLSGSLALLADAGHMLTDTAALLFALLAVQFSRRPPTIRHTFGWLRLTTLA 90

Qy        184 AMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRH 243
              | :: : : ::   :::||::|       : | :|:  |  |:  |::  :||
Db         91 AFVNAIALVVITILIVWEAIER-FRTPRPVEGGMMMAIAVAGLLANILSFWLL------- 142

Qy        244 SHSHSLPSNSPTRGSGCERNHGQD--SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEY 301
                              :|| :  :| |||| :| ||||: ||| :||| || :   :
Db        143 ------------------HHGSEEKNLNVRAAALHVLGDLLGSVGAIIAALIIIW-TGW 182

Qy        302 KIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK-IEDVYSVEDL 360
              ||||  : : |||  : :|:: |:|  :||| |  |::  :|  : : | :| :|  :
Db        183 TPADPILSILVSLLVLRSAWRLLKDSVNELLEGAPVSLDIAELKRRMCREIPEVRNVHHV 242

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              ::|  :   |    :|:|:||      : :  :  | |:: : :   |||::
Db        243 HVW-MVGEKPVMTLHVQVIPPHDH--DALLDQIQHYLMDHYQIEHATIQME 290


Search completed: July 29, 2004, 17:44:50
Job time : 18 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 29, 2004, 17:42:05 ; Search time 19 Seconds
(without alignments)
1165.660 Million cell updates/sec

Title: US-10-024-652-2570
Perfect score: 2183
Sequence: 1 MAGSGAWKRLKSMLRKDDAP.....LQSYRQEVDRTCANCQSSSP 429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	718.5	32.9	372	4	US-10-162-012-40	Sequence 40, Appl
2	555	25.4	398	3	US-09-461-474-17	Sequence 17, Appl
3	549	25.1	448	3	US-09-461-474-8	Sequence 8, Appli
4	536.5	24.6	474	3	US-09-461-474-10	Sequence 10, Appl
5	484.5	22.2	349	3	US-09-461-474-12	Sequence 12, Appl
6	447	20.5	322	4	US-10-162-012-42	Sequence 42, Appl
7	418.5	19.2	359	4	US-09-134-001C-5618	Sequence 5618, Ap
8	372.5	17.1	342	4	US-09-134-001C-4190	Sequence 4190, Ap
9	371	17.0	323	4	US-09-328-352-6181	Sequence 6181, Ap
10	369.5	16.9	320	4	US-09-489-039A-13157	Sequence 13157, A
11	368	16.9	520	4	US-09-252-991A-17058	Sequence 17058, A
12	344.5	15.8	308	4	US-09-543-681A-5795	Sequence 5795, Ap
13	295	13.5	504	4	US-09-461-325-332	Sequence 332, App
14	295	13.5	504	4	US-10-012-542-332	Sequence 332, App
15	288	13.2	363	4	US-09-328-352-4930	Sequence 4930, Ap
16	285	13.1	490	4	US-09-461-325-158	Sequence 158, App
17	285	13.1	490	4	US-10-012-542-158	Sequence 158, App
18	285	13.1	507	4	US-09-795-927-10	Sequence 10, Appl
19	250.5	11.5	157	3	US-09-461-474-14	Sequence 14, Appl
20	209.5	9.6	307	4	US-09-461-325-331	Sequence 331, App
21	209.5	9.6	307	4	US-10-012-542-331	Sequence 331, App
22	199.5	9.1	293	4	US-09-461-325-233	Sequence 233, App
23	199.5	9.1	293	4	US-10-012-542-233	Sequence 233, App
24	188.5	8.6	167	4	US-09-134-001C-2867	Sequence 2867, Ap
25	177.5	8.1	210	4	US-09-252-991A-30886	Sequence 30886, A
26	164.5	7.5	348	4	US-09-107-532A-5255	Sequence 5255, Ap
27	140	6.4	165	4	US-08-858-207A-520	Sequence 520, App
28	134	6.1	316	4	US-09-543-681A-5023	Sequence 5023, Ap
29	126	5.8	372	4	US-09-252-991A-25117	Sequence 25117, A
30	123	5.6	266	4	US-09-540-236-3257	Sequence 3257, Ap
31	121	5.5	443	4	US-09-328-352-8124	Sequence 8124, Ap
32	117.5	5.4	302	4	US-09-252-991A-22091	Sequence 22091, A
33	117.5	5.4	343	4	US-09-328-352-5999	Sequence 5999, Ap
34	117.5	5.4	549	2	US-08-676-279-59	Sequence 59, Appl
35	110	5.0	493	4	US-09-543-681A-4617	Sequence 4617, Ap
36	107.5	4.9	154	4	US-09-134-001C-3700	Sequence 3700, Ap
37	107	4.9	567	4	US-09-711-164-409	Sequence 409, App
38	102	4.7	311	4	US-09-134-001C-5171	Sequence 5171, Ap
39	101.5	4.6	128	4	US-09-328-352-5908	Sequence 5908, Ap
40	99	4.5	325	4	US-09-489-039A-12247	Sequence 12247, A
41	99	4.5	359	1	US-08-153-848-24	Sequence 24, Appl
42	99	4.5	359	3	US-09-299-843A-24	Sequence 24, Appl
43	99	4.5	359	4	US-09-088-337B-24	Sequence 24, Appl
44	99	4.5	359	5	PCT-US93-11153-24	Sequence 24, Appl
45	99	4.5	378	3	US-09-299-843A-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1
US-10-162-012-40
; Sequence 40, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-190001
; CURRENT APPLICATION NUMBER: US/10/162,012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11

```
;  PRIOR APPLICATION NUMBER: US (not assigned)
;  PRIOR FILING DATE: 2002-05-13
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
;   LENGTH: 372
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-162-012-40

  Query Match             32.9%;  Score 718.5;  DB 4;  Length 372;
  Best Local Similarity   38.8%;  Pred. No. 9.1e-73;
  Matches  150;  Conservative   88;  Mismatches  118;  Indels   31;  Gaps    7;

Qy         56 EAPERP-VNGAHPTLQADDDSLLDQD---LPLTNSQLSLK-----VDSCDNCSKQRE--- 103
Db          2 EAKEKQHLLDTRPAIRSYTGSLWQEGAGWIPLPRPGLDLQAIELAAQSNHHCHAQKGPDS 61

Qy        104 --ILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLAL 161
Db         62 HCDPKKGKAQRQLYVASAICLLFMIGEVVGGYLAHSLAVMTDAAHLLTDFASMLISLFSL 121

Qy        162 WLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLIT 221
Db        122 WMSSRPATKTMNFGWQRAEILGALVSVLSIWVVTGVLVYLAVERLISGDYEIDGGTMLIT 181

Qy        222 AAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGD 281
Db        182 SGCAVAVNIIMGLTLHQSGHGHSHGTT-------------NQQEENPSVRAAFIHVIGD 227

Qy        282 LVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNV 341
Db        228 FMQSMGVLVAAYILYFKPEYKYVDPICTFVFSILVLGTTLTILRDVILVLMEGTPKGVDF 287

Qy        342 DYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTF 401
Db        288 TAVRDLLLSVEGVEALHSLHIWALTVAQPVLSVHIAIAQNTDA--QAVLKTASSRLQGKF 345

Qy        402 GMYRCTIQLQSYRQEVDRTCANCQSSS 428
Db        346 HFHTVTIQIEDYSEDM-KDCQACQGPS 371


RESULT 2
US-09-461-474-17
; Sequence 17, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Sakai, Hajime
;  TITLE OF INVENTION: Plant Metal Transporters
;  FILE REFERENCE: BB1303 US NA
;  CURRENT APPLICATION NUMBER: US/09/461,474
;  CURRENT FILING DATE: 1999-12-14
;  EARLIER APPLICATION NUMBER: 60/112,562
;  EARLIER FILING DATE: 1998-12-16
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
;   LENGTH: 398
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-461-474-17

  Query Match             25.4%;  Score 555;  DB 3;  Length 398;
  Best Local Similarity   33.8%;  Pred. No. 4e-54;
  Matches  116;  Conservative   74;  Mismatches  107;  Indels   46;  Gaps    4;

Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRF 172
Db         57 KLCIAVVLCLVFMSVEVVGGIKANSLAILTDAAHLLSDVAAFAISLFSLWAAGWEATPRQ 116

Qy        173 TFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIM 232
Db        117 TYGFFRIEILGALVSIQLIWLLTGILVYEAIIRIVTETSEVNGFLMFLVAAFGLVVNIIM 176

Qy        233 GFLLNQS-------GHRHSHSHSLPSNSPT---------RGSGCERNHGQD--------- 267
Db        177 AVLLGHDHGHSHGHGHGHGHDHHNHSHGVTVTTHHHHHDHEHGHSHGHGEDKHHAHGDVT 236

Qy        268 ------------------SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPIC 308
Db        237 EQLLDKSKTQVAAKEKRKRNINLQGAYLHVLGDSIQSVGVMIGGAIIWYNPEWKIVDLIC 296

Qy        309 TYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSG 368
Db        297 TLAFSVIVLGTTINMIRNILEVLMESTPREIDATKLEKGLLEMEEVVAVHELHIWAITVG 356

Qy        369 KSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
Db        357 KVLLACHVNIRPEADA--DMVLNKVIDYIRREYNISHVTIQIE 397


RESULT 3
US-09-461-474-8
; Sequence 8, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Sakai, Hajime
;  TITLE OF INVENTION: Plant Metal Transporters
;  FILE REFERENCE: BB1303 US NA
;  CURRENT APPLICATION NUMBER: US/09/461,474
;  CURRENT FILING DATE: 1999-12-14
;  EARLIER APPLICATION NUMBER: 60/112,562
;  EARLIER FILING DATE: 1998-12-16
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
;   LENGTH: 448
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-461-474-8

  Query Match             25.1%;  Score 549;  DB 3;  Length 448;
  Best Local Similarity   29.6%;  Pred. No. 2.4e-53;
  Matches  133;  Conservative   90;  Mismatches  146;  Indels   80;  Gaps    9;

Qy         32 SDEAGDE---GLSRFNKLRVVVADDGSEAPERPVNGAHPTLQADDDSLLDQDLPLTNSQL 88
Db          9 SRELGSESWVGLELENRLRRSTQKDSMESH----NPSHSQI-----AEVTMDIAASASGA 59

Qy         89 S----LKVDSCD--NCSKQREILKQRKVKAR-LTIAAVLYLLFMIGELVGGYIANSLAIM 141
Db         60 AGSKFCKGAACDFSDASNSSKDARERSASMRKLIVAVVLCVVFMAVEVVGGIKANSLAIL 119

Qy        142 TDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYE 201
Db        120 TDAAHLLSDVAAFAISLFSLWAAGWEATPRQSYGFFRVEILGALVSIQLIWLLAGILVYE 179

Qy        202 AVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQS-------GHRHSHSH------- 247
Db        180 AVVRLVGESGDVRGSLMFAVSAFGLAVNVLMAVLLGHDHGHGHGHGHGHSHDHGHGHGDS 239

Qy        248 ------------------------------------------SLPSNSPTRGSGCER 262
Db        240 DDGHSHHDDEEQEQGRVHHHEHGHGGAITVTTHHHHHHHHDHDVEEALIKHEGTQSAGRAG 299

Qy        263 NHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFR 322
Db        300 KKPRRNINVHSAYLHVLGDSVQSVGVMVGGAIIWYKPEWKVIDLICTLVFSVVVLFTTIR 359

Qy        323 IIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGS 382
```

```
Db        360 MLRSILEVLMESTPREIDATRLESGLCGMEGVVAVHELHIWAITVGKVLLACHVTIARDA 419

Qy        383 SSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              :  :|:  |    :   : :   |||::
Db        420 DA--DEILDKVIGYIKTEYNISHVTIQVE 446


RESULT 4
US-09-461-474-10
; Sequence 10, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Sakai, Hajime
;  TITLE OF INVENTION: Plant Metal Transporters
;  FILE REFERENCE: BB1303 US NA
;  CURRENT APPLICATION NUMBER: US/09/461,474
;  CURRENT FILING DATE: 1999-12-14
;  EARLIER APPLICATION NUMBER: 60/112,562
;  EARLIER FILING DATE: 1998-12-16
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 10
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: Oryza sativa
US-09-461-474-10

  Query Match             24.6%;  Score 536.5;  DB 3;  Length 474;
  Best Local Similarity   28.5%;  Pred. No. 6.9e-52;
  Matches  109;  Conservative   89;  Mismatches  119;  Indels   65;  Gaps    3;

Qy         93 DSCDNCSKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLS 152
              |  |: :  ::  ::     :| || :| ::||  |:|||  ||||||:||| |:|:|::
Db         93 DFSDSSNSSKDARERMASMRKLIIAVILCIIFMAVEVVGGIKANSLAILTDAAHLLSDVA 152

Qy        153 AIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYE 212
              |  ::| :|| :    | : ::|| |:|:| |::|: |:::| | |:|||: | |: : |
Db        153 AFAISLFSLWAAGWEATPQQSYGFFRIEILGALVSIQLIWLLAGILVYEAIVRLINESGE 212

Qy        213 INGDIMLITAAVGVAVNVIMGFLLNQS---GHRHSHSHSLPSNSPTRGSGCERNHGQD-- 267
              : | :|   :| |: ||:||  ||       || | | |    :    ||  : :| :|
Db        213 VQGSLMFAVSAFGLFVNIIMAVLLGHDHGHGHGHGHGHSHDHDHGGSDHDHHHHEDQE 272

Qy        268 ------------------------------------------------------------SL 269
                                                                           ::
Db        273 HGHVHHHEDGHGNSITVNLHHHPGTGHHHHDAEEPLLKSDAGCDSTQSGAKDAKKARRNI 332

Qy        270 AVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVV 329
              | :|::| ||| :||:||:|   || :|||:|| | ||| :||::| ||| ::: : :
Db        333 NVHSAYLHVLGDSIQSIGVMIGGAIIWYKPEWKIIDLICTLIFSVIVLFTTIKMLRNILE 392

Qy        330 IILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEV 389
              :::|  |  ::  ::  |  :: | :| :|:||::| ||     |: :  : :  :::
Db        393 VLMESTPREIDATSLENGLRDMDGVVAVHELHIWAITVGKVLLACHVTITQDADA--DQM 450

Qy        390 QSKANHLLLNTFGMYRCTIQLQ 411
                |    : : : :   |||::
Db        451 LDKVIGYIKSEYNISHVTIQIE 472


RESULT 5
US-09-461-474-12
; Sequence 12, Application US/09461474
; Patent No. 6278042
; GENERAL INFORMATION:
;  APPLICANT: Allen, Steve
;  APPLICANT: Rafalski, Antoni
;  APPLICANT: Sakai, Hajime
;  TITLE OF INVENTION: Plant Metal Transporters
;  FILE REFERENCE: BB1303 US NA
;  CURRENT APPLICATION NUMBER: US/09/461,474
;  CURRENT FILING DATE: 1999-12-14
;  EARLIER APPLICATION NUMBER: 60/112,562
;  EARLIER FILING DATE: 1998-12-16
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
;   LENGTH: 349
;   TYPE: PRT
;   ORGANISM: Glycine max
US-09-461-474-12

  Query Match             22.2%;  Score 484.5;  DB 3;  Length 349;
  Best Local Similarity   28.6%;  Pred. No. 3.4e-46;
  Matches   99;  Conservative   80;  Mismatches  106;  Indels   61;  Gaps    4;

Qy        125 MIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSA 184
              |  |:|||  ||||||:||| |:|:|:::  ::| :|| :    | | ::|| |:|:| |
Db          5 MTVEVVGGIKANSLAILTDAAHLLSDVASFAISLFSLWAAGWEATPRQSYGFFRIEILGA 64

Qy        185 MISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLL------NQ 238
              ::|: ::::| | |:|||: | |     ::| :| : :| |: ||:||  ||
Db         65 LVSIQMIWLLAGILVYEAIDRIIAGPKNVDGFLMFLVSAFGLVVNIIMALLLGHDHGHRH 124

Qy        239 SGHRHSHSHSLPSNSP--TRGSGCERNHGQD---------------------------- 267
              :|| | | |    :|   |  : |:  | :|
Db        125 AGHSHGHGHDGHGHSHGFTMSTHCDAKHTKDQHHHTHHHDENHPKDAHHHTDEDHLHHHA 184

Qy        268 ----------------------SLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIAD 305
                                     :: |: |::| ||| :||:||:|:|   :| : | ::| |
Db        185 HKEVTELLLGESKGGTKKKKQWNINVQGAYLHVLGDSIQSIGVMIGGAVIWYNPRWQIVD 244

Qy        306 PICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSL 365
              ||| :||::|  ||  :: : : :::|  |  ::   ::  |: :||| :| :|:||::
Db        245 LICTLIFSVIVMGTTINMLRNILEVLMENTPREIDATKLERGLLDMEDVVAVHELHIWAI 304

Qy        366 TSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              | ||      |:::   : :  : |  |    :   : :   |||::
Db        305 TVGKVLLACHVKIRREADA--DLVLDKVIDYIKRVYNISHVTIQIE 348


RESULT 6
US-10-162-012-42
; Sequence 42, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
;  APPLICANT: Curtis, Rory A.J.
;  APPLICANT:  Silos-Santiago, Inmaculada
;  APPLICANT:  Gu, Wei
;  TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
;  FILE REFERENCE: 10448-190001
;  CURRENT APPLICATION NUMBER: US/10/162,012
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: US 60/209,845
;  PRIOR FILING DATE: 2000-06-06
;  PRIOR APPLICATION NUMBER: US 09/875,321
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: PCT/US01/18340
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/209,257
;  PRIOR FILING DATE: 2000-06-05
;  PRIOR APPLICATION NUMBER: US 09/875,423
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: PCT/US01/18398
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/209,238
;  PRIOR FILING DATE: 2000-06-05
;  PRIOR APPLICATION NUMBER: US 09/875,363
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: PCT/US01/18247
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/227,068
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: US 09/928,530
```

```
;  PRIOR FILING DATE: 2001-08-13
;  PRIOR APPLICATION NUMBER: PCT/US01/25475
;  PRIOR FILING DATE: 2001-08-15
;  PRIOR APPLICATION NUMBER: US 60/226,770
;  PRIOR FILING DATE: 2000-08-21
;  PRIOR APPLICATION NUMBER: US 09/934,421
;  PRIOR FILING DATE: 2001-08-21
;  PRIOR APPLICATION NUMBER: PCT/US01/26096
;  PRIOR FILING DATE: 2001-08-21
;  PRIOR APPLICATION NUMBER: US 60/279,281
;  PRIOR FILING DATE: 2001-03-28
;  PRIOR APPLICATION NUMBER: US 10/109,029
;  PRIOR FILING DATE: 2002-03-28
;  PRIOR APPLICATION NUMBER: PCT/US02/09728
;  PRIOR FILING DATE: 2002-03-28
;  PRIOR APPLICATION NUMBER: US 60/290,288
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US (not assigned)
;  PRIOR FILING DATE: 2002-05-13
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 42
;   LENGTH: 322
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: consensus sequence
US-10-162-012-42

  Query Match             20.5%;  Score 447;  DB 4;  Length 322;
  Best Local Similarity   33.8%;  Pred. No. 5.4e-42;
  Matches  112;  Conservative   75;  Mismatches  102;  Indels   42;  Gaps    8;

Qy        116 IAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFG 175
              |:  | || |: :|:|| :: |||:: |||| |:|::: :::|:|| |: | | ::  ||
Db          3 ISLALNLLLMLIKLIGGVLSGSLALLADALHSLSDVASSLISLIALRLAEKPPDEKHPFG 62

Qy        176 FHRLEVLSAMI-SVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLIT------------- 221
               || | |:|:: || || :    :||||::| |  :|||  | :|
Db         63 HHRAETLAALLNSVFLVIVSFLEILYEAIERLISPDYEIPPDAVLAADIMEPEEPGLFEV 122

Qy        222 --AAVGVAVN-----VIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAA 274
                |:|||:      |::| ::| : | :           |  | ::   : :| ||||
Db        123 GGVALGVALGGTALVVLLGLVVNLALHGY----------LRRVG-KKLKSEHNLNVRAA 170

Qy        275 FVHALGDLVQSVGVLIAAYIIRFK-------PEYKIADPICTYVFSLLVAFTTFRIIWDT 327
               :| ||| : |||||||| :| |          :  |||| : : ||:: :| ||:: ::
Db        171 ALHVLGDALSSVGVLIAALLIYFTGYSFKGWKWWYYADPIASILISLIILYTAFRLLKES 230

Qy        328 VVIILEGVPSHLNVD-YIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSK- 385
              |:|:||| ||  :::  ||: |: |  |  | ||:|| | | |  | ||:::      |
Db        231 VLILLEGTPSKEDLERKIKKTLLSIPGVKGVHDLHIWYLGSNKFIASVHVEVDDNLDLKE 290

Qy        386 WEEVQSKANHLLLNTFGMYRCTIQLQSYRQE 416
                :: ::    :|: ||:   |: ::   :|
Db        291 AHDILAEIEREILHKFGIEHVTVHVEPASEE 321


RESULT 7
US-09-134-001C-5618
; Sequence 5618, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5618
;   LENGTH: 359
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5618

  Query Match             19.2%;  Score 418.5;  DB 4;  Length 359;
  Best Local Similarity   27.8%;  Pred. No. 1.1e-38;
  Matches   85;  Conservative   87;  Mismatches  109;  Indels   25;  Gaps    3;

Qy        107 QRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSK 166
              |   |  | :: |: ::| : | ||| ::|||| :::|: |||:|: |: |:::|:: |||
Db         63 QSSSKITLWLSLVITMIFTVVEFVGGIVSNSLALLSDSFHMLSDVLALGLSMVAIYFSSK 122

Qy        167 SPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGV 226
               ||| :|:|| |||:: | :: | : ::   ::|| : | ||    :   ||:: | :|:
Db        123 PPTKNYTYGFLRLEIIVAFLNGLALIVISLGIMYEGIMRIIHPR-PVESGIMILIAFIGL 181

Qy        227 AVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSV 286
                |::: :|  |  :                          :::: :::|  | :|||: |:
Db        182 IANIVLTIILMISLKK----------------------ENNINIQSALWHFIGDLLNSL 218

Qy        287 GVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKE 346
              |::|  :| |   : | ||| : : ||::   ::|| :   :::| ||   : | |
Db        219 GIIVAFVLIHF-TGWNIVDPIISILISLIILRGGYKIIKNASKVLMERVPDRYDTDEIMG 277

Qy        347 ALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRC 406
              |:  :| |  : : ::||:|: :|:   |: |             :| : ||   :|:
Db        278 AMKDVEGVIDIHEFHLWSVTTNQSSLSAHVVLSDDYIKSPYATINKVSDLLKTQYGLEHV 337

Qy        407 TIQLQS 412
              |:|:::
Db        338 TLQIEN 343


RESULT 8
US-09-134-001C-4190
; Sequence 4190, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
;  TITLE OF INVENTION:  EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC-007
;  CURRENT APPLICATION NUMBER: US/09/134,001C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/064,964
;  PRIOR FILING DATE: 1997-11-08
;  PRIOR APPLICATION NUMBER: US 60/055,779
;  PRIOR FILING DATE: 1997-08-14
;  NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4190
;   LENGTH: 342
;   TYPE: PRT
;   ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4190

  Query Match             17.1%;  Score 372.5;  DB 4;  Length 342;
  Best Local Similarity   26.9%;  Pred. No. 1.8e-33;
  Matches   86;  Conservative   88;  Mismatches  119;  Indels   27;  Gaps    5;

Qy        111 KARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTK 170
              |  |  : ::  |||: |::||::|||||:::|  |||:|  :: : |:|   : |  ||
Db         47 KKVLMFSFIIISLFMLVEIIGGFLANSLALLSDGFHMLSDAISLGVALIAFIYAEKHATK 106

Qy        171 RFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNV 230
                |:|: | |:|:|: : : ::|:   :: ||::|  :    |:    | | : :|: ||:
Db        107 SKTYGYKRFEILAALFNGVTLFIISIIIIEAIRRFLEPP-EVQSKEMFIISVIGLMVNI 165

Qy        231 IMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLI 290
              |:  |: : |                  : :|   :| :| ||:|  ||||   ||| ::
Db        166 IVAILMFKGG------------------DTSH---NLNMRGAFLHVLGDLFGSVGAIV 202
```

```
Qy        291 AAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK 350
              |: :| :   : ||||| : : ||::  : : |   :: |::|| |: :::: : :|: |
Db        203 ASLLI-WGFNFTIADPIASILVSLIILKSAYGISKSSLNILMEGTPNDIDLNAVIKAISK 261

Qy        351 IEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQL 410
               | : :| | ::|:::: :    |  :    | :  |   |:    |    :   ||||
Db        262 DERIQNVHDCHVWTISNDMNALSCHAVVPEYLSVQTCETMLKSIESDLLQLNIQHMTIQL 321

Qy        411 QS--YRQEVDRTCANCQSSS 428
              ::  :: :    |:     |
Db        322 ETPEHKHDESTLCSGIHEHS 341


RESULT 9
US-09-328-352-6181
; Sequence 6181, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6181
;   LENGTH: 323
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-6181

  Query Match             17.0%;  Score 371;  DB 4;  Length 323;
  Best Local Similarity   29.4%;  Pred. No. 2.4e-33;
  Matches   92;  Conservative   68;  Mismatches  117;  Indels   36;  Gaps    5;

Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRF 172
              :|||| |   |:| |:: | |  ||||:::|| || || :|: : |:|: :| :    :
Db         26 KLTIALALTTTFLIVEVIAGLITQSLALLSDAAHMFTDAAALAIALVAIQISKRPADNKR 85

Qy        173 TFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIM 232
              |||: | |:|:|: : |:::::  ::|||| |       ||    ||| | :|: :|:|
Db         86 TFGYQRFEILAALFNALMLFVVAIYILYEAYIR-FSQPPEIQSVGMLIVATIGLVINLIS 144

Qy        233 GFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAA 292
              :|  | :                          :|| |: |::  | | : |||||:| |
Db        145 MKILMSSAN------------------------NSLNVKGAYLEVLSDALGSVGVIIGA 179

Qy        293 YIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIE 352
              || |   | | | :   :    |   |: ::  :: |:|||||  :::: ::  |: :
Db        180 IIIYFTNWYWI-DTLIAVLIGFWVLPRTWVLLKQSINILLEGVPEEVDIEKLRADLLSLN 238

Qy        353 DVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQS 412
              | |:  | :|::||      ||:       :|   :   |  :| :  |:   |:|::
Db        239 GVESIHQLKVWAITSKNIHLTVHLFAPEADRNK---LYQDAVEMLSHEHGIGEVTLQIE- 294

Qy        413 YRQEVDRTCANCQ 425
                   |   |||
Db        295 -----DDAEINCQ 302


RESULT 10
US-09-489-039A-13157
; Sequence 13157, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 13157
;   LENGTH: 320
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13157

  Query Match             16.9%;  Score 369.5;  DB 4;  Length 320;
  Best Local Similarity   29.6%;  Pred. No. 3.5e-33;
  Matches   99;  Conservative   76;  Mismatches  118;  Indels   41;  Gaps    9;

Qy         79 QDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSL 138
              :|  : :|      | : || : :            ||  | ::   ||: | ||| |: ||
Db          3 RDDTMAHSHSHSSVQAPDNSNAR----------RLLWAFIVTAGFMLIEAVGGAISGSL 51

Qy        139 AIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFL 198
              |:: || ||||| :|::  |||:  :|: |  | |||: ||  |:| :: : : ::   :
Db         52 ALLADAGHMLTDSAALLFALLAVRFASRPPNARHTFGWLRLTTLAAFLNAIALVVITMLI 111

Qy        199 LYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGS 258
              ::||:||  |    : |  |:: |  |:  ||:  ::|   || |
Db        112 VWEAIQRFQHPQ-PVAGVTMMVIAVAGLLANVLAFWIL----HRGSE------------- 153

Qy        259 GCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAF 318
                |||     | |||| :| |||: ||| ::|| :|     :   ||| : : | ||
Db        154 --ERN-----LNVRAAALHVLGDLLGSVGAIVAAVVI-LTTGWTPVDPILSVLVSCLVLR 205

Qy        319 TTFRIIWDTVVIILEGVPSHLNVDYIKEALMK-IEDVYSVEDLNIWSLTSGKSTAIVHIQ 377
              : :|:: :::  :||| |  |:|  ::  : : | :|  |  :::| |   |    :|:|
Db        206 SAWRLLKESLNELLEGAPRSLDVAALQRDIRRSIPEVRDVHHVHVW-LVGEKPVMTLHVQ 264

Qy        378 LIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQ 411
              ::|      : : ::  | | : : :   |:|::
Db        265 VVPPHDH--DALLNRILHFLEHKYEIAHATVQME 296


RESULT 11
US-09-252-991A-17058
; Sequence 17058, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT:  Marc J. Rubenfield et al.
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION:  AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE:  1999-02-18
;  PRIOR APPLICATION NUMBER:  US 60/074,788
;  PRIOR FILING DATE:  1998-02-18
;  PRIOR APPLICATION NUMBER:  US 60/094,190
;  PRIOR FILING DATE:  1998-07-27
;  NUMBER OF SEQ ID NOS:  33142
; SEQ ID NO 17058
;   LENGTH: 520
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17058

  Query Match             16.9%;  Score 368;  DB 4;  Length 520;
  Best Local Similarity   29.3%;  Pred. No. 1.2e-32;
  Matches   93;  Conservative   72;  Mismatches  116;  Indels   36;  Gaps    6;

Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRF 172
              ||  | :|   |:: |:||| :  |||:::|| |||||  |: : | |: :: :    |
Db        237 RLKWALLLTGSFLVAEVVGGILTGSLALISDAAHMLTDAVALAIALAAINIARRPTNDRL 296

Qy        173 TFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIM 232
              |:|:|| |:|:|  :  |:: :  ::|| | :| ::   ||    ||: | :|: ||:
Db        297 TYGYHRFEILAAAFNAFLLFGVAFYILYAAYER-LNQPAEIQSVGMLVIAVLGLLVNLAS 355

Qy        233 GFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAA 292
                ||              :| :|        :|| |: |::    |:: |:||::||
```

```
Db        356 MRLL----------------APAQG---------NSLNVKGAYLEVWSDMLGSLGVIVAA 390

Qy        293 YIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIE 352
              :|||    :   | :   :    |    |: :: ::: ::|||||  : :  ::|||: |
Db        391 IVIRF-TGWAWVDSLVAVLIGFWVLPRTWILLCESLHVLLEGVPKEIQLAELREALLGIP 449

Qy        353 DVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQS 412
              |  : ||::||:|||| :   |:   |       | :      || : : :   |:||:
Db        450 GVTGLHDLHVWSITSGKISLTGHLVYDPALVDA-EALLGTVKALLHDRYEIEHSTLQLE- 507

Qy        413 YRQEVDRTCANCQSSSP 429
                  | |  |:  |
Db        508 -------TSACAQAEEP 517


RESULT 12
US-09-543-681A-5795
; Sequence 5795, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5795
;   LENGTH: 308
;   TYPE: PRT
;   ORGANISM: Proteus mirabilis
US-09-543-681A-5795

  Query Match            15.8%;  Score 344.5;  DB 4;  Length 308;
  Best Local Similarity  29.1%;  Pred. No. 2.3e-30;
  Matches   88;  Conservative   67;  Mismatches  118;  Indels   29;  Gaps    4;

Qy        111 KARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTK 170
              ::|| ||  |    ||: | :|||:  |||:::||:||:||  |::| |:|:    |:
Db         25 ESRLLIAFALTGSFMVIEFIGGYLTESLALISDAMHMMTDAFALLLALIAIHAGRKAADL 84

Qy        171 RFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNV 230
                |:|: | |:|:|  :: |::  :  ::|||| || :    :|    ||: |  |: :|
Db         85 FRTYGYARFEILAAAMNALILMAVAFYILYEAYQR-LSSPPDIQSMGMLVVAIAGLIINF 143

Qy        231 IMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLI 290
              |  :|  :                        :|| |: |::    |:: ||||:|
Db        144 ISMKMLTSA------------------------KDESLNVKGAYLEVWADMLGSVGVII 178

Qy        291 AAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMK 350
              | ||| :   ::: | |   :   :|   |: :: : : |:|||||  :|:  : : :
Db        179 GAVII-WLTGWQLVDSIIAVLIGFMVFPRTWILLKECLNILLEGVPKGINLTEVTDTINN 237

Qy        351 IEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQL 410
              | |  ||::|:||  |      ||   |   |  | ::   :  |   | :   |:|:
Db        238 TAGVASSHDLHLWALTQSKLILSGHIVYQPNVDS--ETLRLAIDKQLREQFHINHITLQM 295

Qy        411 QS 412
              :|
Db        296 ES 297


RESULT 13
US-09-461-325-332
; Sequence 332, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 94 Human Secreted Proteins
;  FILE REFERENCE: PZ029P1
;  CURRENT APPLICATION NUMBER: US/09/461,325A
;  CURRENT FILING DATE: 1999-12-14
;  EARLIER APPLICATION NUMBER: PCT/US99/13418
;  EARLIER FILING DATE: 1999-06-15
;  EARLIER APPLICATION NUMBER: 60/089,507
;  EARLIER FILING DATE: 1998-06-16
;  EARLIER APPLICATION NUMBER: 60/089,508
;  EARLIER FILING DATE: 1998-06-16
;  EARLIER APPLICATION NUMBER: 60/089,509
;  EARLIER FILING DATE: 1998-06-16
;  EARLIER APPLICATION NUMBER: 60/089,510
;  EARLIER FILING DATE: 1998-06-16
;  EARLIER APPLICATION NUMBER: 60/090,112
;  EARLIER FILING DATE: 1998-06-22
;  EARLIER APPLICATION NUMBER: 60/090,113
;  EARLIER FILING DATE: 1998-06-22
;  NUMBER OF SEQ ID NOS: 532
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 332
;   LENGTH: 504
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (148)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (403)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-461-325-332

  Query Match            13.5%;  Score 295;  DB 4;  Length 504;
  Best Local Similarity  23.4%;  Pred. No. 2.3e-24;
  Matches   99;  Conservative   76;  Mismatches  124;  Indels  124;  Gaps   13;

Qy        105 LKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLS 164
              ::  : : ||      |  :||: |:|   : :|||:::|: |||:|: |:::  |:|   :
Db          3 IRHERNRGRLLCMLALTFMFMVLEVVVSRVTSSLAMLSDSFHMLSDVLALVVALVAERFA 62

Qy        165 SKS-PTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAA 223
              ::  |:: |||: | ||: |::: : :   |   :| ||::| |  :      ::|
Db         63 RRTHATQKNTFGWIRAEVMGALVNAIFLTGLCFAILLEAIERFIEPHEMQQPLVVLGVGV 122

Qy        224 VGVAVNVI--------MGFLLNQSGHRHS-----HSHSLP-------------------- 250
              |: |||:          ||    ||| ||      | | ||
Db        123 AGLLVNVLGLCLFHHHSGF-SQDSGHXHSHGGHGHGHGLPKGPRVKSTRPGSSDINVAPG 181

Qy        251 ----------------SNS------------PTRGSGC----------ERNH-------- 264
                             |||            |  |              | :|
Db        182 EQGPDQEETNTLVANTSNSNGLKLDPADPENPRSGDTVEVQVNGNLVREPDHMELEEDRA 241

Qy        265 GQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRF-------------------KPEYKI- 303
              ||  | :|  |:| ||| : || |:: | : |                    ||  :|
Db        242 GQ--LNMRGVFLHVLGDALGSVIVVVNALVFYFSWKGCSEGDFCVNPCFPDPCKPFVEII 299

Qy        304 -----------------ADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKE 346
                              ||     |   :: :||: :: :: :|:|: ||  :::  : :
Db        300 NSTHASVYEAGPCWVLYLDPTLCVVMVCILLYTTYPLLKESALILLQTVPKQIDIRNLIK 359

Qy        347 ALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRC 406
              |  :| |  | :|::| |   :  |  ||:    :|  : || :|    : :  |::
Db        360 ELRNVEGVEEVHELHVWQLAGSRIIATAHIKCEDPTS--YMEV-AKXIKDVFHNHGIHAT 416

Qy        407 TIQ 409
              |||
Db        417 TIQ 419


RESULT 14
US-10-012-542-332
; Sequence 332, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
```

```
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 94 Human Secreted Proteins
;  FILE REFERENCE: PZ029P1
;  CURRENT APPLICATION NUMBER: US/10/012,542
;  CURRENT FILING DATE:  2001-12-12
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
;  PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
;  NUMBER OF SEQ ID NOS: 532
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 332
;   LENGTH: 504
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (148)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (403)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-012-542-332

  Query Match             13.5%;  Score 295;  DB 4;  Length 504;
  Best Local Similarity   23.4%;  Pred. No. 2.3e-24;
  Matches   99;  Conservative   76;  Mismatches  124;  Indels  124;  Gaps   13;

Qy        105 LKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLS 164
              ::  : : ||      |  :||: |:|   : :|||:::|: |||:|: |::: |:|   :
Db          3 IRHERNRGRLLCMLALTFMFMVLEVVVSRVTSSLAMLSDSFHMLSDVLALVVALVAERFA 62

Qy        165 SKS-PTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAA 223
               ::  |:: |||: | ||: |::: : :  |   :| ||::| |  :      ::|
Db         63 RRTHATQKNTFGWIRAEVMGALVNAIFLTGLCFAILLEAIERFIEPHEMQQPLVVLGVGV 122

Qy        224 VGVAVNVI---------MGFLLNQSGHRHS-----HSHSLP-------------------- 250
               |: |||:          ||    ||| ||      | | ||
Db        123 AGLLVNVLGLCLFHHHSGF-SQDSGHXHSHGGHGHGHGLPKGPRVKSTRPGSSDINVAPG 181

Qy        251 ----------------SNS------------PTRGSGC----------ERNH-------- 264
                             |||            |  |             | :|
Db        182 EQGPDQEETNTLVANTSNSNGLKLDPADPENPRSGDTVEVQVNGNLVREPDHMELEEDRA 241

Qy        265 GQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRF-------------------KPEYKI- 303
              ||  | :|  |:| ||| : || |:: | :  |                  ||  :|
Db        242 GQ--LNMRGVFLHVLGDALGSVIVVVNALVFYFSWKGCSEGDFCVNPCFPDPCKPFVEII 299

Qy        304 ----------------ADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKE 346
                             ||     |   :: :||: :: :: :|:|: ||  :::  : :
Db        300 NSTHASVYEAGPCWVLYLDPTLCVVMVCILLYTTYPLLKESALILLQTVPKQIDIRNLIK 359

Qy        347 ALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRC 406
               |  :| |  | :|::| |   :  |  ||:    :|  : || :|    : :  |::
Db        360 ELRNVEGVEEVHELHVWQLAGSRIIATAHIKCEDPTS--YMEV-AKXIKDVFHNHGIHAT 416

Qy        407 TIQ 409
              |||
Db        417 TIQ 419


RESULT 15
US-09-328-352-4930
```

```
; Sequence 4930, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4930
;   LENGTH: 363
;   TYPE: PRT
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-4930

  Query Match             13.2%;  Score 288;  DB 4;  Length 363;
  Best Local Similarity   27.1%;  Pred. No. 8.1e-24;
  Matches   73;  Conservative   66;  Mismatches  120;  Indels   10;  Gaps    5;

Qy        113 RLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSK-SPTKR 171
              |: || :|    |: |: ||:  ||:|:: |  || : : |: |   |   :   |   |
Db         59 RILIATILTASMMVLEVFGGWFFNSMALLADGWHMSSHMLALGLAYFAYRAARHYSNDHR 118

Qy        172 FTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVI 231
              |:||  ::|:|:|   | :|: ::  |: ::::||  :   ||  :  :  | :|: :|:|
Db        119 FSFGTWKIEILAGYSSAILLMVVAIFMAFQSIQRLFN-PVEIFYNEAIPIAILGLVINLI 177

Qy        232 MGFLLNQSG---HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGV 288
               :||:  |   | | | |     :    | |    |    | :|||:| : | | ||
Db        178 CAWLLHDDGHHHHHHHHHHHHHHHHHHEHGHG----HHHHDLNQKAAFLHVVADAVTSVFA 233

Qy        289 LIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEAL 348
              ::| :  ::   :   | |   : ::|||  :| :: :|   :|:    |  || |:| :
Db        234 IVALFAGKYF-GWDFLDAILGILGAILVAKWSFGLMKETGKTLLDAEMDHPVVDEIREVI 292

Qy        349 MKIEDVYSVEDLNIWSLTSGKSTAIVHIQ 377
               :      : |:::|  :  || : |: ::
Db        293 AEFPKHVEITDIHVWKVAKGKFSCILALE 321


Search completed: July 29, 2004, 17:45:20
Job time : 20 secs
```

Blank Sheet

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 29, 2004, 17:37:14 ; Search time 54 Seconds
(without alignments)
2244.683 Million cell updates/sec

Title: US-10-024-652-2570
Perfect score: 2183
Sequence: 1 MAGSGAWKRLKSMLRKDDAP.....LQSYRQEVDRTCANCQSSSP 429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2183	100.0	429	5	ABG61811	Abg61811 Prostate
2	2183	100.0	429	5	ABJ17930	Abj17930 Zinc tran
3	2183	100.0	429	5	ABJ17931	Abj17931 Zinc tran
4	2183	100.0	429	5	ABJ17933	Abj17933 Zinc tran
5	2183	100.0	438	5	AAE24062	Aae24062 Human pro
6	2179	99.8	429	5	ABJ17932	Abj17932 Zinc tran
7	2179	99.8	429	5	ABJ17934	Abj17934 Zinc tran
8	2179	99.8	429	7	AAE38600	Aae38600 Human zin
9	2166.5	99.2	430	7	ADB82764	Adb82764 Human pro
10	747	34.2	358	5	ABB83085	Abb83085 Transport
11	739	33.9	326	4	ABU52726	Abu52726 Human int
12	723	33.1	356	6	ABU54543	Abu54543 Human NOV
13	722.5	33.1	372	5	AAE22905	Aae22905 Human tra
14	722.5	33.1	372	5	ABB83084	Abb83084 Human tra
15	718.5	32.9	372	6	ABB99884	Abb99884 Human cat
16	679	31.1	323	6	ABU54544	Abu54544 Human NOV
17	653.5	29.9	369	5	ABG60224	Abg60224 Human Zin
18	652.5	29.9	472	4	ABB60324	Abb60324 Drosophil
19	648.5	29.7	369	5	ABG60226	Abg60226 Human Zin
20	648	29.7	320	4	AAB60094	Aab60094 Human tra
21	648	29.7	320	5	AAU99907	Aau99907 Human 842
22	648	29.7	320	5	AAE17562	Aae17562 Human pan
23	638	29.2	320	4	ABU52725	Abu52725 Human int
24	617.5	28.3	422	5	ABG60225	Abg60225 Human Zin
25	616	28.2	388	7	ADC77681	Adc77681 Human 443
26	599.5	27.5	1677	4	ABB69976	Abb69976 Drosophil
27	555	25.4	344	3	AAG43479	Aag43479 Arabidops
28	555	25.4	344	3	AAG22264	Aag22264 Arabidops
29	555	25.4	398	3	AAG43478	Aag43478 Arabidops
30	555	25.4	398	3	AAG22263	Aag22263 Arabidops
31	549	25.1	448	4	AAE09321	Aae09321 Maize zin
32	538.5	24.7	322	5	ABP69136	Abp69136 Human pol
33	536.5	24.6	474	4	AAE09322	Aae09322 Rice zinc
34	534	24.5	375	3	AAG31822	Aag31822 Arabidops
35	534	24.5	385	3	AAG31821	Aag31821 Arabidops
36	532	24.4	452	4	ABB65511	Abb65511 Drosophil
37	530	24.3	274	7	ADB65759	Adb65759 Human pro
38	525	24.0	330	3	AAG22265	Aag22265 Arabidops
39	525	24.0	330	3	AAG43480	Aag43480 Arabidops
40	524.5	24.0	359	3	AAG31823	Aag31823 Arabidops
41	523.5	24.0	420	6	ABP72086	Abp72086 Perennial
42	484.5	22.2	349	4	AAE09323	Aae09323 Soybean z
43	447	20.5	322	5	AAU99928	Aau99928 Human 842
44	447	20.5	322	5	AAU99931	Aau99931 Human 850
45	439.5	20.1	319	6	ABM72087	Abm72087 Staphyloc

ALIGNMENTS

RESULT 1
ABG61811
ID ABG61811 standard; protein; 429 AA.
XX
AC ABG61811;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #12.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
DR WPI; 2002-471335/50.
DR N-PSDB; ABK92126.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
PS Claim 27; Page 310; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
SQ Sequence 429 AA;

Query Match 100.0%; Score 2183; DB 5; Length 429;
Best Local Similarity 100.0%; Pred. No. 2.1e-234;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

RESULT 2
ABJ17930
ID ABJ17930 standard; protein; 429 AA.
XX
AC ABJ17930;
XX
DT 16-JAN-2003 (first entry)
XX
DE Zinc transporter protein 108P5H8 v-1.
XX
KW Cytostatic; gene therapy; vaccine; zinc transporter protein 108P5H8;
KW cancer; breast; colon; ovarian; lung; humoral; cellular immune response;
KW passive immunisation.
XX
OS Unidentified.
XX
PN WO200260953-A2.
XX
PD 08-AUG-2002.
XX
PF 17-DEC-2001; 2001WO-US049133.
XX
PR 15-DEC-2000; 2000US-0256210P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Challita-Eid PM, Faris M, Afar DEH, Hubert RS, Mitchell SC;
PI Levin E, Morrison KJM, Raitano AB, Jakobovits A;
XX
DR WPI; 2002-627469/67.
DR N-PSDB; AAL53482.
XX
PT Composition comprising a substance that modulates the status of a zinc
PT transporter protein (108P5H8), useful in diagnosing and treating patients
PT with cancer that express 108P5H8, such as breast, colon, ovarian or lung
PT cancer.
XX
PS Claim 33; Fig 2A; 309pp; English.
XX
CC The invention relates to a new composition comprising a substance that
CC modulates the status of a zinc transporter protein, designated as
CC 108P5H8, or a molecule that is modulated by 108P5H8. The composition is
CC useful in diagnosing, preventing, prognosticating or treating patients
CC with cancer that expresses 108P5H8, such as breast, colon, ovarian or
CC lung cancer. The 108P5H8 gene or its fragment can be used to elicit a
CC humoral or cellular immune response. The antibodies are useful in active
CC or passive immunisation. The 108P5H8 polynucleotides are useful as probes
CC and primers for the amplification or detection of 108P5H8 genes, as
CC coding sequences for directing the expression of 108P5H8 polypeptides, or
CC as tools for modulating or inhibiting the expression of 108P5H8 genes.
CC The polynucleotides of the invention can be used to treat disorders by
CC gene therapy. This sequence represents a zinc transporter protein 108P5H8
CC sequence of the invention
XX
SQ Sequence 429 AA;

Query Match 100.0%; Score 2183; DB 5; Length 429;
Best Local Similarity 100.0%; Pred. No. 2.1e-234;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

RESULT 3

ABJ17931
ID ABJ17931 standard; protein; 429 AA.
XX
AC ABJ17931;
XX
DT 16-JAN-2003 (first entry)
XX
DE Zinc transporter protein 108P5H8 v-2.
XX
KW Cytostatic; gene therapy; vaccine; zinc transporter protein 108P5H8;
KW cancer; breast; colon; ovarian; lung; humoral; cellular immune response;
KW passive immunisation.
XX
OS Unidentified.
XX
PN WO200260953-A2.
XX
PD 08-AUG-2002.
XX
PF 17-DEC-2001; 2001WO-US049133.
XX
PR 15-DEC-2000; 2000US-0256210P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Challita-Eid PM, Faris M, Afar DEH, Hubert RS, Mitchell SC;
PI Levin E, Morrison KJM, Raitano AB, Jakobovits A;
XX
DR WPI; 2002-627469/67.
DR N-PSDB; AAL53483.
XX
PT Composition comprising a substance that modulates the status of a zinc
PT transporter protein (108P5H8), useful in diagnosing and treating patients
PT with cancer that express 108P5H8, such as breast, colon, ovarian or lung
PT cancer.
XX
PS Claim 33; Fig 2B; 309pp; English.
XX
CC The invention relates to a new composition comprising a substance that
CC modulates the status of a zinc transporter protein, designated as
CC 108P5H8, or a molecule that is modulated by 108P5H8. The composition is
CC useful in diagnosing, preventing, prognosticating or treating patients
CC with cancer that expresses 108P5H8, such as breast, colon, ovarian or
CC lung cancer. The 108P5H8 gene or its fragment can be used to elicit a
CC humoral or cellular immune response. The antibodies are useful in active
CC or passive immunisation. The 108P5H8 polynucleotides are useful as probes
CC and primers for the amplification or detection of 108P5H8 genes, as
CC coding sequences for directing the expression of 108P5H8 polypeptides, or
CC as tools for modulating or inhibiting the expression of 108P5H8 genes.
CC The polynucleotides of the invention can be used to treat disorders by
CC gene therapy. This sequence represents a zinc transporter protein 108P5H8
CC sequence of the invention
XX
SQ Sequence 429 AA;

Query Match 100.0%; Score 2183; DB 5; Length 429;
Best Local Similarity 100.0%; Pred. No. 2.1e-234;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

RESULT 4
ABJ17933
ID ABJ17933 standard; protein; 429 AA.
XX
AC ABJ17933;
XX
DT 16-JAN-2003 (first entry)
XX
DE Zinc transporter protein 108P5H8 v-1 and v-2.
XX
KW Cytostatic; gene therapy; vaccine; zinc transporter protein 108P5H8;
KW cancer; breast; colon; ovarian; lung; humoral; cellular immune response;
KW passive immunisation.
XX
OS Unidentified.
XX
PN WO200260953-A2.
XX
PD 08-AUG-2002.
XX
PF 17-DEC-2001; 2001WO-US049133.
XX
PR 15-DEC-2000; 2000US-0256210P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Challita-Eid PM, Faris M, Afar DEH, Hubert RS, Mitchell SC;
PI Levin E, Morrison KJM, Raitano AB, Jakobovits A;
XX
DR WPI; 2002-627469/67.
XX
PT Composition comprising a substance that modulates the status of a zinc
PT transporter protein (108P5H8), useful in diagnosing and treating patients
PT with cancer that express 108P5H8, such as breast, colon, ovarian or lung
PT cancer.
XX
PS Disclosure; Fig 3A; 309pp; English.
XX
CC The invention relates to a new composition comprising a substance that
CC modulates the status of a zinc transporter protein, designated as
CC 108P5H8, or a molecule that is modulated by 108P5H8. The composition is
CC useful in diagnosing, preventing, prognosticating or treating patients
CC with cancer that expresses 108P5H8, such as breast, colon, ovarian or
CC lung cancer. The 108P5H8 gene or its fragment can be used to elicit a
CC humoral or cellular immune response. The antibodies are useful in active
CC or passive immunisation. The 108P5H8 polynucleotides are useful as probes
CC and primers for the amplification or detection of 108P5H8 genes, as
CC coding sequences for directing the expression of 108P5H8 polypeptides, or
CC as tools for modulating or inhibiting the expression of 108P5H8 genes.
CC The polynucleotides of the invention can be used to treat disorders by
CC gene therapy. This sequence represents a zinc transporter protein 108P5H8
CC sequence of the invention
XX
SQ Sequence 429 AA;

```
  Query Match             100.0%;  Score 2183;  DB 5;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-234;
  Matches  429;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 5
AAE24062
ID   AAE24062 standard; protein; 438 AA.
XX
AC   AAE24062;
XX
DT   23-SEP-2002  (first entry)
XX
DE   Human prostate specific protein (PSP) #5.
XX
KW   Human; prostate specific protein; PSP; prostate specific nucleic acid;
KW   vaccine; transgenic; prostate cancer; gene therapy; transgenic animal;
KW   PSNA.
XX
OS   Homo sapiens.
XX
PN   WO200224718-A1.
XX
PD   28-MAR-2002.
XX
PF   19-SEP-2001; 2001WO-US029386.
XX
PR   19-SEP-2000; 2000US-0233746P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Sun Y,  Recipon H,  Cafferkey R,  Ali S;
XX
DR   WPI; 2002-471216/50.
XX
PT   Novel isolated prostate specific polypeptide useful for identifying,
PT   diagnosing, monitoring, staging, imaging, and treating prostate cancer
PT   and non-cancerous disease states in prostate.
XX
PS   Claim 37; Page 198-199; 210pp; English.
XX
CC   The invention relates to prostate specific proteins (PSP) and prostate
CC   specific nucleic acids (PSNA). Sequences of the invention are useful for
CC   identifying, diagnosing, monitoring, staging, imaging and treating
CC   prostate cancer and non-cancerous disease states in prostate. They are
CC   also useful for producing engineered prostate tissue for treatment and
CC   research. The PSNA sequences are used in gene therapy and for producing
CC   transgenic animals and cells. The invention is also used as vaccines. The
CC   present sequence is human prostate specific protein of the invention
XX
SQ   Sequence 438 AA;

  Query Match             100.0%;  Score 2183;  DB 5;  Length 438;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-234;
  Matches  429;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 6
ABJ17932
ID   ABJ17932 standard; protein; 429 AA.
XX
AC   ABJ17932;
XX
DT   16-JAN-2003  (first entry)
XX
DE   Zinc transporter protein 108P5H8 v-3.
XX
KW   Cytostatic; gene therapy; vaccine; zinc transporter protein 108P5H8;
KW   cancer; breast; colon; ovarian; lung; humoral; cellular immune response;
KW   passive immunisation.
XX
OS   Unidentified.
XX
PN   WO200260953-A2.
XX
PD   08-AUG-2002.
XX
PF   17-DEC-2001; 2001WO-US049133.
XX
PR   15-DEC-2000; 2000US-0256210P.
XX
```

```
PA   (AGEN-) AGENSYS INC.
XX
PI   Challita-Eid PM,  Faris M,  Afar DEH,  Hubert RS,  Mitchell SC;
PI   Levin E,  Morrison KJM,  Raitano AB,  Jakobovits A;
XX
DR   WPI; 2002-627469/67.
DR   N-PSDB; AAL53484.
XX
PT   Composition comprising a substance that modulates the status of a zinc
PT   transporter protein (108P5H8), useful in diagnosing and treating patients
PT   with cancer that express 108P5H8, such as breast, colon, ovarian or lung
PT   cancer.
XX
PS   Claim 33; Fig 2C; 309pp; English.
XX
CC   The invention relates to a new composition comprising a substance that
CC   modulates the status of a zinc transporter protein, designated as
CC   108P5H8, or a molecule that is modulated by 108P5H8. The composition is
CC   useful in diagnosing, preventing, prognosticating or treating patients
CC   with cancer that expresses 108P5H8, such as breast, colon, ovarian or
CC   lung cancer. The 108P5H8 gene or its fragment can be used to elicit a
CC   humoral or cellular immune response. The antibodies are useful in active
CC   or passive immunisation. The 108P5H8 polynucleotides are useful as probes
CC   and primers for the amplification or detection of 108P5H8 genes, as
CC   coding sequences for directing the expression of 108P5H8 polypeptides, or
CC   as tools for modulating or inhibiting the expression of 108P5H8 genes.
CC   The polynucleotides of the invention can be used to treat disorders by
CC   gene therapy. This sequence represents a zinc transporter protein 108P5H8
CC   sequence of the invention
XX
SQ   Sequence 429 AA;

  Query Match             99.8%;  Score 2179;  DB 5;  Length 429;
  Best Local Similarity   99.8%;  Pred. No. 5.9e-234;
  Matches  428;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFEFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429


RESULT 7
ABJ17934
ID   ABJ17934 standard; protein; 429 AA.
XX
AC   ABJ17934;
XX
DT   16-JAN-2003  (first entry)
XX
DE   Zinc transporter protein 108P5H8 v-3 #2.
XX
KW   Cytostatic; gene therapy; vaccine; zinc transporter protein 108P5H8;
KW   cancer; breast; colon; ovarian; lung; humoral; cellular immune response;
KW   passive immunisation.
XX
OS   Unidentified.
XX
PN   WO200260953-A2.
XX
PD   08-AUG-2002.
XX
PF   17-DEC-2001; 2001WO-US049133.
XX
PR   15-DEC-2000; 2000US-0256210P.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Challita-Eid PM,  Faris M,  Afar DEH,  Hubert RS,  Mitchell SC;
PI   Levin E,  Morrison KJM,  Raitano AB,  Jakobovits A;
XX
DR   WPI; 2002-627469/67.
XX
PT   Composition comprising a substance that modulates the status of a zinc
PT   transporter protein (108P5H8), useful in diagnosing and treating patients
PT   with cancer that express 108P5H8, such as breast, colon, ovarian or lung
PT   cancer.
XX
PS   Disclosure; Fig 3B; 309pp; English.
XX
CC   The invention relates to a new composition comprising a substance that
CC   modulates the status of a zinc transporter protein, designated as
CC   108P5H8, or a molecule that is modulated by 108P5H8. The composition is
CC   useful in diagnosing, preventing, prognosticating or treating patients
CC   with cancer that expresses 108P5H8, such as breast, colon, ovarian or
CC   lung cancer. The 108P5H8 gene or its fragment can be used to elicit a
CC   humoral or cellular immune response. The antibodies are useful in active
CC   or passive immunisation. The 108P5H8 polynucleotides are useful as probes
CC   and primers for the amplification or detection of 108P5H8 genes, as
CC   coding sequences for directing the expression of 108P5H8 polypeptides, or
CC   as tools for modulating or inhibiting the expression of 108P5H8 genes.
CC   The polynucleotides of the invention can be used to treat disorders by
CC   gene therapy. This sequence represents a zinc transporter protein 108P5H8
CC   sequence of the invention
XX
SQ   Sequence 429 AA;

  Query Match             99.8%;  Score 2179;  DB 5;  Length 429;
  Best Local Similarity   99.8%;  Pred. No. 5.9e-234;
  Matches  428;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFEFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

RESULT 8
AAE38600
ID AAE38600 standard; protein; 429 AA.
XX
AC AAE38600;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human zinc transporter 4 protien.
XX
KW Human; urological disorder; urinary incontinence; gene therapy; cancer;
KW kidney disorder; overactive; oversensitive bladder; dysfunction; bladder;
KW urethra; overflow urinary incontinence; stress urinary incontinence;
KW nervous system; prostatitis; benign prostatic hyperplasia; nephrotrophic;
KW zinc transporter 4.
XX
OS Homo sapiens.
XX
PN WO2003061573-A2.
XX
PD 31-JUL-2003.
XX
PF 16-JAN-2003; 2003WO-US001450.
XX
PR 18-JAN-2002; 2002US-0349511P.
PR 28-FEB-2002; 2002US-0360500P.
PR 15-MAR-2002; 2002US-0365041P.
PR 19-APR-2002; 2002US-0374063P.
PR 14-AUG-2002; 2002US-0403468P.
PR 27-SEP-2002; 2002US-0414262P.
PR 21-OCT-2002; 2002US-0419986P.
PR 05-NOV-2002; 2002US-0423809P.
PR 26-NOV-2002; 2002US-0429797P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Silos-Santiago I, Karicheti V;
XX
DR WPI; 2003-598705/56.
DR N-PSDB; AAD58501.
XX
PT Identifying a compound for treating urological disorders, for example
PT urinary incontinence by assaying the ability of the compound to modulate
PT the nucleic acid expression or polypeptide activity.
XX
PS Disclosure; Page 217-218; 0pp; English.
XX
CC The present relates to a method for identifying a compound for treating
CC urological disorders e.g., urinary incontinence including overactive/
CC oversensitive bladder, overflow urinary incontinence, stress urinary
CC incontinence caused by dysfunction of the bladder, urethra or central or
CC peripheral nervous system, prostatitis, benign prostatic hyperplasia,
CC cancer of the prostate or kidney disorders. The method is also useful for
CC modulating hyperplasia in a cell and treating a subject having a
CC urological disorder. The invention is also used in gene therapy. The
CC present sequence is human zinc transporter 4 protein
XX
SQ Sequence 429 AA;

```
  Query Match           99.8%;  Score 2179;  DB 7;  Length 429;
  Best Local Similarity  99.8%;  Pred. No. 5.9e-234;
  Matches  428;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFEFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLE 180

Qy        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSG 240

Qy        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 HRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPE 300

Qy        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 YKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDL 360

Qy        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 NIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRT 420

Qy        421 CANCQSSSP 429
              |||||||||
Db        421 CANCQSSSP 429
```

RESULT 9
ADB82764
ID ADB82764 standard; protein; 430 AA.
XX
AC ADB82764;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human protein sequence useful for the treatment of cancer (SeqID 1545).
XX
KW human; prostate; cancer; cytostatic; gene therapy; vaccine;
KW immune response.
XX
OS Homo sapiens.
XX
PN WO2003050236-A2.
XX
PD 19-JUN-2003.
XX
PF 04-SEP-2002; 2002WO-US028214.
XX
PR 07-DEC-2001; 2001US-00012697.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Kassam A, Lamson G, Drmanac R;
PI Crkvenjakov R, Dickson M, Drmanac S, Labat I, Leshkowitz D, Kita D;
PI Garcia V, Jones LW, Stache-Crain B, Scott EM;
XX
DR WPI; 2003-513972/48.
XX
PT New polynucleotides derived from human prostate, useful for modulating
PT immune response to prevent or treat cancer.
XX

```
PS   Claim 18; SEQ ID NO 1545; 188pp; English.
XX
CC   This invention relates to novel isolated polynucleotides of human origin,
CC   particularly isolated from the human prostate. Specifically, it refers to
CC   the diagnostics and therapeutics comprising these novel human
CC   polynucleotides, and includes the derived probes, antisense
CC   oligonucleotides and antibodies thereof. The identification of these
CC   human prostate genes that can inhibit tumour growth is useful for
CC   understanding the progression and nature of complex diseases such as
CC   cancer, and hence they are important in the drug discovery process. The
CC   present invention describes these polynucleotides and encoded
CC   polypeptides as exhibiting cytostatic activity, and through gene therapy
CC   and/ or vaccines they can be used to modulate the immune response for the
CC   prevention or treatment of cancers, particularly of the prostate, but
CC   also for breast, lung and colon cancer. This polypeptide sequence is a
CC   human protein sequence useful for the treatment of cancer, used in an
CC   exemplification of the invention. NOTE: These sequences are not given in
CC   the specification but are provided on the WIPO website.
XX
SQ   Sequence 430 AA;

  Query Match             99.2%;  Score 2166.5;  DB 7;  Length 430;
  Best Local Similarity   99.5%;  Pred. No. 1.5e-232;
  Matches  428;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MAGSGAWKRLKSMLRKDDAPLFLNDTSAFDFSDEAGDEGLSRFNKLRVVVADDGSEAPER 60

Qy         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 PVNGAHPTLQADDDSLLDQDLPLTNSQLSLKVDSCDNCSKQREILKQRKVKARLTIAAVL 120

Qy        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRL- 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        121 YLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLG 180

Qy        180 EVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQS 239
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 XVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQS 240

Qy        240 GHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKP 299
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKP 300

Qy        300 EYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVED 359
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVED 360

Qy        360 LNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDR 419
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDR 420

Qy        420 TCANCQSSSP 429
              ||||||||||
Db        421 TCANCQSSSP 430


RESULT 10
ABB83085
ID   ABB83085 standard; protein; 358 AA.
XX
AC   ABB83085;
XX
DT   02-OCT-2002  (first entry)
XX
DE   Transporter related protein.
XX
KW   Rat; transporter protein; zinc transporter; pharmacogenomic analysis;
KW   diagnosis; drug screening; gene therapy.
XX
OS   Rattus norvegicus.
XX
PN   WO200224910-A2.
XX
PD   28-MAR-2002.
XX
PF   20-SEP-2001; 2001WO-US029218.
XX
PR   20-SEP-2000; 2000US-0234160P.
PR   19-OCT-2000; 2000US-00691219.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Wei M,  Ketchum KA,  Di Francesco V,  Beasley EM;
XX
DR   WPI; 2002-404954/43.
XX
PT   Novel human transporter proteins, related to zinc transporter subfamily,
PT   useful as model for developing human therapeutic targets and serves as
PT   target for human therapeutics.
XX
PS   Disclosure; Page 75; 75pp; English.
XX
CC   The invention relates to an isolated human transporter protein that is
CC   related to the zinc transporter subfamily. Polynucleotides and
CC   polypeptides of the invention are useful for treating a disease or
CC   condition mediated by human transporter protein. The proteins also
CC   provide a target for diagnosing a disease or predisposition to disease
CC   mediated by the peptide, and in pharmacogenomic analysis. The peptides
CC   are also useful for treating a disorders characterised by absence of,
CC   inappropriate or unwanted expression of the protein. The nucleic acids
CC   are also useful in drug screening assays and as a target for treatment by
CC   the compounds identified through drug screening. The invention also
CC   provides vectors for gene therapy in patients with aberrant expression of
CC   the gene encoding the transporter protein. The gene of the invention has
CC   been found to be expressed in humans in the kidney, testis, heart,
CC   placenta, small intestine and liver. The current sequence represents a
CC   rat amino acid sequence that is relative to the transporter protein of
CC   the invention
XX
SQ   Sequence 358 AA;

  Query Match             34.2%;  Score 747;  DB 5;  Length 358;
  Best Local Similarity   42.0%;  Pred. No. 5.2e-74;
  Matches  148;  Conservative   78;  Mismatches  108;  Indels   18;  Gaps    4;

Qy         82 PLTNSQLSLKVDSCDNCSKQREI-----LKQRKVKARLTIAAVLYLLFMIGELVGGYIAN 136
              |:    : | | |    |  |::       :::: : :| :|: : |:|||||::|||:|
Db         19 PVNLPSVELAVQSNHYCHAQKDSGSHPNSEKQRARRKLYVASAICLVFMIGEIIGGYLAQ 78

Qy        137 SLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMG 196
              |||||||  |:||| ::::::| :||:||:   ||   ||: | |:| |::||| :::: |
Db         79 SLAIMTDAAHLLTDFASMLISLFSLWVSSRPATKTMNFGWQRAEILGALLSVLSIWVVTG 138

Qy        197 FLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTR 256
              |:| |||| |  :||| || ||||:   ||||:|||  |:|||| ||| ||   :|
Db        139 VLVYLAVQRLISGDYEIKGDTMLITSGCAVAVNIIMGLALHQSGHGHSHGHSHEDSS--- 195

Qy        257 GSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLV 316
                    | : :|||||:| :|||:||||||:||||| ||||||  |||||::||:||
Db        196 -------QQQQNPSVRAAFIHVVGDLLQSVGVLVAAYIIYFKPEYKYVDPICTFLFSILV 248

Qy        317 AFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHI 376
              ||  |: | :::::|| |  ::   :|  |: :: | ::  |:||:||  :    |||
Db        249 LGTTLTILRDVILVLMEGTPKGVDFTTVKNLLLSVDGVEALHSLHIWALTVAQPVLSVHI 308

Qy        377 QLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQSSS 428
              :  :  : | | |  |  :  |||::|| ::: ::|  || |
Db        309 AIAQNVDA--QAVLKVARDRLQGKFNFHTMTIQIESYSEDM-KSCQECQGPS 357


RESULT 11
ABU52726
ID   ABU52726 standard; protein; 326 AA.
XX
```

```
AC   ABU52726;
XX
DT   14-APR-2003  (first entry)
XX
DE   Human intracellular transport and trafficking DKFZphfbr2_62f10 homologue.
XX
KW   Human; gene therapy; vaccine; disease treatment; detection.
XX
OS   Homo sapiens.
XX
PN   WO200112659-A2.
XX
PD   22-FEB-2001.
XX
PF   18-AUG-2000; 2000WO-IB001496.
XX
PR   18-AUG-1999;   99US-0149499P.
PR   28-SEP-1999;   99US-0156503P.
XX
PA   (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI   Wiemann S;
XX
DR   WPI; 2001-327840/34.
XX
PT   Nucleic acids having the sequences of clones isolated from libraries of
PT   different human tissues, useful in recombinant DNA methodologies.
XX
PS   Example III; Page 266; 1095pp; English.
XX
CC   This invention describes novel polynucleotides and polypeptides isolated
CC   from human cDNA libraries which can be used for gene therapy or in
CC   vaccines. The polynucleotides of the invention and antibodies encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate polypeptide expression. The products of the
CC   invention may also be used to identify modulators of expression and
CC   activity and to down regulate expression and activity. The antibodies of
CC   the invention may also be used as diagnostic agents for detecting the
CC   presence of polypeptides in samples. This sequence represents a homologue
CC   of a polypeptide described in the disclosure of the invention
XX
SQ   Sequence 326 AA;

  Query Match             33.9%;  Score 739;  DB 4;  Length 326;
  Best Local Similarity   44.0%;  Pred. No. 3.5e-73;
  Matches  146;  Conservative   70;  Mismatches   96;  Indels   20;  Gaps    4;

Qy          97 NCSKQREILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIIL 156
               |  |||         : :| :|: : |:|||||::|||:| |||||||| |:||| ::::: 
Db          14 NSEKQR-------ARRKLYVASAICLVFMIGEIIGGYLAQSLAIMTDAAHLLTDFASMLI 66

Qy         157 TLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGD 216
               :| :||:||:  ||   ||: | |:| |::||| :::: | |:| |||| |  :||| || 
Db          67 SLFSLWVSSRPATKTMNFGWQRAEILGALLSVLSIWVVTGVLVYLAVQRLISGDYEIKGD 126

Qy         217 IMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFV 276
                ||||:   ||||:|||  |:|||| ||| ||   :|           | : :|||||: 
Db         127 TMLITSGCAVAVNIIMGLALHQSGHGHSHGHSHEDSS----------QQQQNPSVRAAFI 176

Qy         277 HALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVP 336
               | :|||:||||||:||||| ||||||  |||||::||:||  ||  |: | :::::|| | 
Db         177 HVVGDLLQSVGVLVAAYIIYFKPEYKYVDPICTFLFSILVLGTTLTILRDVILVLMEGTP 236

Qy         337 SHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHL 396
                 ::   :|  |: :: | ::  |:||:||  :    ||| :     :  : |   |  
Db         237 KGVDFTTVKNLLLSVDGVEALHSLHIWALTVAQPVLSVHIAIAQNVDA--QAVLKVARDR 294

Qy         397 LLNTFGMYRCTIQLQSYRQEVDRTCANCQSSS 428
               |   |  :  |||::|| ::: ::|  ||  |
Db         295 LQGKFNFHTMTIQIESYSEDM-KSCQECQGPS 325


RESULT 12
ABU54543
ID   ABU54543 standard; protein; 356 AA.
XX
AC   ABU54543;
XX
DT   03-JUN-2003  (first entry)
XX
DE   Human NOVX polypeptide #2.
XX
KW   Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW   hypertension; congenital heart defect; aortic stenosis; valve disease;
KW   atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW   pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW   tuberous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW   obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease; immune disorder; haematopoietic disorder;
KW   haemophilia; hypercoagulation; Crohn's disease; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200281498-A2.
XX
PD   17-OCT-2002.
XX
PF   03-APR-2002; 2002WO-US010780.
XX
PR   03-APR-2001; 2001US-0281086P.
PR   03-APR-2001; 2001US-0281136P.
PR   05-APR-2001; 2001US-0281863P.
PR   05-APR-2001; 2001US-0281906P.
PR   06-APR-2001; 2001US-0282020P.
PR   10-APR-2001; 2001US-0282930P.
PR   10-APR-2001; 2001US-0282934P.
PR   12-APR-2001; 2001US-0283512P.
PR   13-APR-2001; 2001US-0283710P.
PR   17-APR-2001; 2001US-0284234P.
PR   19-APR-2001; 2001US-0285325P.
PR   20-APR-2001; 2001US-0285381P.
PR   20-APR-2001; 2001US-0285609P.
PR   23-APR-2001; 2001US-0285748P.
PR   23-APR-2001; 2001US-0285890P.
PR   24-APR-2001; 2001US-0286068P.
PR   25-APR-2001; 2001US-0286292P.
PR   27-APR-2001; 2001US-0287213P.
PR   02-MAY-2001; 2001US-0288257P.
PR   29-MAY-2001; 2001US-0294164P.
PR   30-MAY-2001; 2001US-0294484P.
PR   18-JUN-2001; 2001US-0298952P.
PR   19-JUN-2001; 2001US-0299237P.
PR   19-JUN-2001; 2001US-0299276P.
PR   12-SEP-2001; 2001US-0318750P.
PR   25-SEP-2001; 2001US-0324800P.
PR   25-SEP-2001; 2001US-0324802P.
PR   27-SEP-2001; 2001US-0325684P.
PR   17-OCT-2001; 2001US-0330143P.
PR   14-NOV-2001; 2001US-0332131P.
PR   14-NOV-2001; 2001US-0332240P.
PR   14-NOV-2001; 2001US-0332779P.
PR   21-NOV-2001; 2001US-0332115P.
PR   04-DEC-2001; 2001US-0337621P.
PR   03-JAN-2002; 2002US-0345783P.
PR   16-JAN-2002; 2002US-0350251P.
PR   02-APR-2002; 2002US-00114270.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Guo X,  Kekuda R,  Miller CE,  Malyankar UM,  Spytek KA;
PI   Patturajan M,  Liu X,  Gusev VY,  Li L,  Vernet CAM,  Zerhusen BD;
PI   Gorman L,  Shenoy SG,  Pena CEA,  Smithson G,  Burgess CE,  Gerlach V;
PI   Padigaru M,  Shimkets RA,  Gangolli EA,  Taupier RJ,  Casman SJ,  Ji W;
PI   Anderson DW,  Leite MW,  Rastelli L,  Edinger SR,  Stone DJ;
PI   Macdougall JR,  Rothenberg ME,  Mazur A,  Millet I,  Peyman JA;
PI   Ellerman K;
```

XX
DR WPI; 2003-046858/04.
DR N-PSDB; ABX72171.
XX
PT New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
PS Claim 1; Page 94; 666pp; English.
XX
CC The invention relates to human polypeptides, termed NOVX, and the
CC polynucleotides encoding them. The polypeptides and polynucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic
CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuberous sclerosis, scleroderma,
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABU54542-ABU54647 represent human NOVX polypeptides
CC of the invention
XX
SQ Sequence 356 AA;

Query Match 33.1%; Score 723; DB 6; Length 356;
Best Local Similarity 40.5%; Pred. No. 2.5e-71;
Matches 143; Conservative 82; Mismatches 106; Indels 22; Gaps 4;

Qy 81 LPLTNSQLSLKVDSCDNCSKQRE-----ILKQRKVKARLTIAAVLYLLFMIGELVGGYIA 135
Db 20 LPLDLQAIELAAQSNHHCHAQKGPDSHCDPKKGKAQRQLYVASAICLLFMIGEVVGGYLA 79

Qy 136 NSLAIMTDALHMLTDLSAIILTLLALWLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILM 195
Db 80 HSLAVMTDAAHLLTDFASMLISLFSLWMSSRPATKTMNFGWQRAEILGALVSVLSIWVVT 139

Qy 196 GFLLYEAVQRTIHMNYEINGDIMLITAAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPT 255
Db 140 GVLVYLAVERLISGDYEIDGGTMLITSGCAVAVNIMMGLTLHQSGHGHSHGTT------- 192

Qy 256 RGSGCERNHGQDSLAVRAAFVHALGDLVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLL 315
Db 193 -------NQQEENPSVRAAFIHVIGDFMQSMGVLVAAYILYFKPEYKYVDPICTFVFSIL 245

Qy 316 VAFTTFRIIWDTVVIILEGVPSHLNVDYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVH 375
Db 246 VLGTTLTILRDVILVLMEGTPKGVDFTAVRDLLLSVEGVEALHSLHIWALTVAQPVLSVH 305

Qy 376 IQLIPGSSSKWEEVQSKANHLLLNTFGMYRCTIQLQSYRQEVDRTCANCQSSS 428
Db 306 IAIAQNTDA--QAVLKTASSRLQGKFHFHTVTIQIEDYSEDM-KDCQACQGPS 355

RESULT 13
AAE22905
ID AAE22905 standard; protein; 372 AA.
XX
AC AAE22905;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human transporter and ion channel (TRICH) 4.
XX
KW Human; transporter and ion channel; TRICH; transport disorder;
KW diabetes mellitus; angina; Alzheimer's disease; neurological; epilepsy;
KW stroke; Huntington's disease; meningitis; muscle; myocarditis; cancer;
KW infectious myositis; arrhythmia; asthma; immunological; gene therapy;
KW acquired immunodeficiency syndrome; AIDS; allergy; atherosclerosis;
KW cell proliferative disorder; cerebroprotective; cirrhosis; hepatitis;
KW transgenic; neuroprotective; anticonvulsant; nootropic; cytostatic;
KW antiinflammatory; hepatotropic; psoriasis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 141. .159
FT /label= Transmembrane_domain
XX
PN WO200222684-A2.
XX
PD 21-MAR-2002.
XX
PF 14-SEP-2001; 2001WO-US028938.
XX
PR 15-SEP-2000; 2000US-0232685P.
PR 22-SEP-2000; 2000US-0234842P.
PR 29-SEP-2000; 2000US-0236882P.
PR 05-OCT-2000; 2000US-0239057P.
PR 13-OCT-2000; 2000US-0240540P.
PR 18-OCT-2000; 2000US-0241700P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lee EA, Yue H, Lal PG, Walia NK, Baughn MR, Warren BA, Lee S;
PI Sanjanwala MS, Yao MG, Ramkumar J, Thornton M, Gandhi AR;
PI Policky JL, Elliott VS, Arvizu C, Raumann BE, Bruns CM, Naini A;
PI Hafalia AJA, Nguyen DB, Xu Y, Lu DAM, Ison CH, Griffin JA;
PI Reddy RM, Burford N;
XX
DR WPI; 2002-393948/42.
DR N-PSDB; AAD36301.
XX
PT Polypeptides of human transporters and ion channels, useful for
PT diagnosing, treating or preventing transport, neurological, muscle,
PT immunological and cell proliferative disorders.
XX
PS Claim 1; Page 144-145; 204pp; English.
XX
CC The invention relates to human transporters and ion channels (TRICH) and
CC their corresponding nucleic acid sequences. TRICH is useful for screening
CC an agonist/antagonist that modulates its activity. TRICH is useful as an
CC immunogen for preparing antibodies which are useful for diagnosing a
CC condition of disease associated with its expression in a subject, and for
CC detecting and purifying it from a sample. TRICH DNA is useful as probe or
CC a primer for assessing toxicity of a test compound. Composition
CC comprising TRICH or its agonist is useful for treating a disease or
CC condition associated with decreased expression of functional TRICH and
CC composition comprising TRICH antagonist is useful for treating a disease
CC or condition associated with TRICH overexpression of TRICH. TRICH
CC sequence is used in the diagnosis and treatment of transport disorder
CC e.g. diabetes mellitus, angina, Alzheimer's disease; neurological
CC disorder e.g. epilepsy, stroke, Huntington's disease, bacterial and viral
CC meningitis, muscle disorder e.g. myocarditis, infectious myositis,
CC arrhythmias, asthma, immunological disorder e.g. acquired
CC immunodeficiency syndrome (AIDS), allergies, atherosclerosis; and cell
CC proliferative disorders e.g. cirrhosis, hepatitis, psoriasis and cancers.
CC TRICH DNA is used in gene therapy. TRICH DNA is useful for creating
CC knockin humanised animals (pigs) or transgenic animals (mice or rats) to
CC model human disease. The present sequence is human TRICH protein
XX
SQ Sequence 372 AA;

Query Match 33.1%; Score 722.5; DB 5; Length 372;
Best Local Similarity 39.0%; Pred. No. 3e-71;
Matches 151; Conservative 88; Mismatches 117; Indels 31; Gaps 7;

Qy 56 EAPERP-VNGAHPTLQADDDSLLDQD---LPLTNSQLSLK-----VDSCDNCSKQRE--- 103
Db 2 EAKEKQHLLDARPAIRSYTGSLWQEGAGWIPLPRPGLDLQAIELAAQSNHHCHAQKGPDS 61

Qy 104 --ILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLAL 161
Db 62 HCDPKKGKAQRQLYVASAICLLFMIGEVVGGYLAHSLAVMTDAAHLLTDFASMLISLFSL 121

```
Qy        162 WLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLIT 221
              |:||:  ||   ||: | |:| |::||| :::: | |:| ||:| |  :|||:|  ||||
Db        122 WMSSRPATKTMNFGWQRAEILGALVSVLSIWVVTGVLVYLAVERLISGDYEIDGGTMLIT 181

Qy        222 AAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGD 281
              :   ||||:|||  |:|||| |||  :              |  ::: :|||||:| :||
Db        182 SGCAVAVNIIMGLTLHQSGHGHSHGTT-------------NQQEENPSVRAAFIHVIGD 227

Qy        282 LVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNV 341
               :||:|||:||||: ||||||  |||||:|||:||  ||  |: | :::::|| |  ::
Db        228 FMQSMGVLVAAYILYFKPEYKYVDPICTFVFSILVLGTTLTILRDVILVLMEGTPKGVDF 287

Qy        342 DYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTF 401
              ::: |: :| | ::  |:||:||  :    ||| :   : :  : |   |:  |   |
Db        288 TAVRDLLLSVEGVEALHSLHIWALTVAQPVLSVHIAIAQNTDA--QAVLKTASSRLQGKF 345

Qy        402 GMYRCTIQLQSYRQEVDRTCANCQSSS 428
              :  |||:: | ::: : |  ||  |
Db        346 HFHTVTIQIEDYSEDM-KDCQACQGPS 371
```

RESULT 14
ABB83084
ID ABB83084 standard; protein; 372 AA.
XX
AC ABB83084;
XX
DT 02-OCT-2002 (first entry)
XX
DE Human transporter protein.
XX
KW Human; transporter protein; zinc transporter; pharmacogenomic analysis;
KW diagnosis; drug screening; gene therapy; kidney; testis; heart; placenta;
KW small intestine; liver; chromosome 1.
XX
OS Homo sapiens.
XX
PN WO200224910-A2.
XX
PD 28-MAR-2002.
XX
PF 20-SEP-2001; 2001WO-US029218.
XX
PR 20-SEP-2000; 2000US-0234160P.
PR 19-OCT-2000; 2000US-00691219.
XX
PA (PEKE) PE CORP NY.
XX
PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-404954/43.
DR N-PSDB; ABN83946, ABN83947.
XX
PT Novel human transporter proteins, related to zinc transporter subfamily,
PT useful as model for developing human therapeutic targets and serves as
PT target for human therapeutics.
XX
PS Claim 1(a); Fig 2; 75pp; English.
XX
CC The invention relates to an isolated human transporter protein that is
CC related to the zinc transporter subfamily. Polynucleotides and
CC polypeptides of the invention are useful for treating a disease or
CC condition mediated by human transporter protein. The proteins also
CC provide a target for diagnosing a disease or predisposition to disease
CC mediated by the peptide, and in pharmacogenomic analysis. The peptides
CC are also useful for treating a disorders characterised by absence of,
CC inappropriate or unwanted expression of the protein. The nucleic acids
CC are also useful in drug screening assays and as a target for treatment by
CC the compounds identified through drug screening. The invention also
CC provides vectors for gene therapy in patients with aberrant expression of
CC the gene encoding the transporter protein. The gene of the invention has
CC been found to be expressed in humans in the kidney, testis, heart,
CC placenta, small intestine and liver. The gene has been localised to human
CC chromosome 1. The current sequence represents the human transporter
CC protein
XX
SQ Sequence 372 AA;

```
  Query Match             33.1%;  Score 722.5;  DB 5;  Length 372;
  Best Local Similarity   39.0%;  Pred. No. 3e-71;
  Matches  151;  Conservative   88;  Mismatches  117;  Indels   31;  Gaps    7;

Qy         56 EAPERP-VNGAHPTLQADDDSLLDQD---LPLTNSQLSLK-----VDSCDNCSKQRE--- 103
              || |:  :  | | :::   ||  :    :||    | |:       |  :|  |:
Db          2 EAKEKQHLLDARPAIRSYTGSLWQEGAGWIPLPRPGLDLQAIELAAQSNHHCHAQKGPDS 61

Qy        104 --ILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLAL 161
                |: | : :| :|: : |||||||:||||:|:|||:|||| |:||| ::::::| :|
Db         62 HCDPKKGKAQRQLYVASAICLLFMIGEVVGGYLAHSLAVMTDAAHLLTDFASMLISLFSL 121

Qy        162 WLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLIT 221
              |:||:  ||   ||: | |:| |::||| :::: | |:| ||:| |  :|||:|  ||||
Db        122 WMSSRPATKTMNFGWQRAEILGALVSVLSIWVVTGVLVYLAVERLISGDYEIDGGTMLIT 181

Qy        222 AAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGD 281
              :   ||||:|||  |:|||| |||  :              |  ::: :|||||:| :||
Db        182 SGCAVAVNIIMGLTLHQSGHGHSHGTT-------------NQQEENPSVRAAFIHVIGD 227

Qy        282 LVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNV 341
               :||:|||:||||: ||||||  |||||:|||:||  ||  |: | :::::|| |  ::
Db        228 FMQSMGVLVAAYILYFKPEYKYVDPICTFVFSILVLGTTLTILRDVILVLMEGTPKGVDF 287

Qy        342 DYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTF 401
              ::: |: :| | ::  |:||:||  :    ||| :   : :  : |   |:  |   |
Db        288 TAVRDLLLSVEGVEALHSLHIWALTVAQPVLSVHIAIAQNTDA--QAVLKTASSRLQGKF 345

Qy        402 GMYRCTIQLQSYRQEVDRTCANCQSSS 428
              :  |||:: | ::: : |  ||  |
Db        346 HFHTVTIQIEDYSEDM-KDCQACQGPS 371
```

RESULT 15
ABB99884
ID ABB99884 standard; protein; 372 AA.
XX
AC ABB99884;
XX
DT 21-FEB-2003 (first entry)
XX
DE Human cation transporter 84226.
XX
KW Human; cation transporter 84226; integral membrane protein;
KW divalent metal cation; metal transport-related disorder;
KW cellular toxicity; pancreatic disorder; metabolic disorder; obesity;
KW diabetes; pain disorder; cellular proliferation disorder; cancer;
KW cellular differentiation disorder; cardiovascular disorder;
KW kidney disorder; pharmacogenomic marker; drug screening;
KW chromosome mapping; tissue typing; forensic biology; cytostatic;
KW analgesic; cardiant; nephrotropic; gene therapy.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT Domain   1. .73
FT   /label= Cytoplasmic_domain_1
FT Modified-site   61. .64
FT   /note= "Ser is O-phosphorylated by casein kinase II"
FT Domain   74. .361
FT   /label= Ca2+_efflux_domain
FT Domain   74. .95
FT   /label= Transmembrane_domain_1
FT Region   96. .106
FT   /label= Extracellular_loop_1
FT Domain   107. .123
FT   /label= Transmembrane_domain_2
FT Domain   124. .140

```
FT                   /label= Cytoplasmic_domain_2
FT   Modified-site   124. .126
FT                   /note= "Ser is O-phosphorylated by protein kinase C"
FT   Domain          141. .163
FT                   /label= Transmembrane_domain_3
FT   Modified-site   143. .148
FT                   /note= "Gly is N-myristoylated"
FT   Region          164. .177
FT                   /label= Extracellular_loop_2
FT   Domain          178. .196
FT                   /label= Transmembrane_domain_4
FT   Modified-site   183. .188
FT                   /note= "Gly is N-myristoylated"
FT   Domain          197. .218
FT                   /label= Cytoplasmic_domain_3
FT                   /note= "Histidine-rich"
FT   Modified-site   199. .202
FT                   /note= "Glycosaminoglycan attachment site"
FT   Modified-site   216. .218
FT                   /note= "Ser is O-phosphorylated by protein kinase C"
FT   Domain          219. .243
FT                   /label= Transmembrane_domain_5
FT   Modified-site   233. .238
FT                   /note= "Gly is N-myristoylated"
FT   Region          244. .252
FT                   /label= Extracellular_loop_3
FT   Domain          253. .277
FT                   /label= Transmembrane_domain_6
FT   Modified-site   264. .269
FT                   /note= "Gly is N-myristoylated"
FT   Domain          278. .372
FT                   /label= Cytoplasmic_domain_4
FT                   /note= "C-terminal tail"
FT   Modified-site   280. .285
FT                   /note= "Gly is N-myristoylated"
FT   Modified-site   281. .283
FT                   /note= "Thr is O-phosphorylated by protein kinase C"
FT   Modified-site   338. .340
FT                   /note= "Ser is O-phosphorylated by protein kinase C"
XX
PN   WO200279427-A2.
XX
PD   10-OCT-2002.
XX
PF   28-MAR-2002; 2002WO-US009728.
XX
PR   28-MAR-2001; 2001US-0279281P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Curtis RAJ;
XX
DR   WPI; 2003-046802/04.
DR   N-PSDB; ABQ79299.
XX
PT   New isolated human cation transporter 84226 polypeptides and nucleic acid
PT   molecules, for treating or preventing a disorder of aberrant activity of
PT   84226-expressing cell, e.g. pancreatic cancer or metal transport-related
PT   disorder.
XX
PS   Claim 5; Page 106-107; 117pp; English.
XX
CC   This sequence represents a novel human cation transporter family member,
CC   designated 84226. Members of this family of proteins are integral
CC   membrane proteins which increase cellular tolerance to divalent metal
CC   cations such as zinc, cadmium and cobalt by mediating their diffusion
CC   across membranes. Cation transporter 84226 shares the structural features
CC   of the cation transporter family, including six transmembrane domains, a
CC   histidine-rich loop between the fourth and fifth transmembrane domains, a
CC   cation efflux domain, and a long C-terminal tail. Expression analysis
CC   indicated that the cation transporter is highly expressed in pancreas and
CC   is slightly expressed in heart, kidney, skeletal muscle and small
CC   intestine. Cation transporter 84226 nucleic acid molecules and proteins
CC   are useful for diagnosing, treating or preventing a disorder
CC   characterised by aberrant activity or expression of cation transporter
CC   84226, such as conditions involving metal transport-related disorders;
CC   disorders associated with cellular toxicity resulting from aberrant or
CC   deficient cation diffusion; pancreatic disorders (e.g., pancreatic cancer
CC   or pancreatitis); metabolic disorders (e.g., obesity, anorexia nervosa,
CC   cachexia, lipid disorders or diabetes); pain disorders; or aberrant or
CC   deficient cellular proliferation or differentiation (e.g., cancers).
CC   Cation transporter 84226 molecules are also useful as diagnostic targets
CC   and therapeutic agents for controlling cardiovascular disorders (e.g.,
CC   hypertension, atherosclerosis, coronary artery disease or arrhythmias),
CC   kidney disorders (e.g., amyloidosis, glomerulonephritis or Goodpasture's
CC   syndrome), or as pharmacogenomic markers. They may additionally be used
CC   in drug screening, and the polynucleotides can be used for chromosome
CC   mapping, tissue typing or in forensic biology
XX
SQ   Sequence 372 AA;
```

```
  Query Match             32.9%;  Score 718.5;  DB 6;  Length 372;
  Best Local Similarity   38.8%;  Pred. No. 8.5e-71;
  Matches  150;  Conservative   88;  Mismatches  118;  Indels   31;  Gaps    7;

Qy         56 EAPERP-VNGAHPTLQADDDSLLDQD---LPLTNSQLSLK-----VDSCDNCSKQRE--- 103
              || |:  :    | :::   ||  :    :||    | |:       |  :|  |:
Db          2 EAKEKQHLLDTRPAIRSYTGSLWQEGAGWIPLPRPGLDLQAIELAAQSNHHCHAQKGPDS 61

Qy        104 --ILKQRKVKARLTIAAVLYLLFMIGELVGGYIANSLAIMTDALHMLTDLSAIILTLLAL 161
                 |: | : :| :|: : |||||||:||||:|:|||:|||| |:||| ::::::| :|
Db         62 HCDPKKGKAQRQLYVASAICLLFMIGEVVGGYLAHSLAVMTDAAHLLTDFASMLISLFSL 121

Qy        162 WLSSKSPTKRFTFGFHRLEVLSAMISVLLVYILMGFLLYEAVQRTIHMNYEINGDIMLIT 221
              |:||:  ||   ||: | |:| |::||| :::: | |:| ||:| |  :|||:|  ||||
Db        122 WMSSRPATKTMNFGWQRAEILGALVSVLSIWVVTGVLVYLAVERLISGDYEIDGGTMLIT 181

Qy        222 AAVGVAVNVIMGFLLNQSGHRHSHSHSLPSNSPTRGSGCERNHGQDSLAVRAAFVHALGD 281
              :   ||||:|||  |:|||| |||  :              |  ::: :|||||:| :||
Db        182 SGCAVAVNIIMGLTLHQSGHGHSHGTT-------------NQQEENPSVRAAFIHVIGD 227

Qy        282 LVQSVGVLIAAYIIRFKPEYKIADPICTYVFSLLVAFTTFRIIWDTVVIILEGVPSHLNV 341
               :||:|||:||||: ||||||  |||||:|||:||  ||  |: | :::::|| |  ::
Db        228 FMQSMGVLVAAYILYFKPEYKYVDPICTFVFSILVLGTTLTILRDVILVLMEGTPKGVDF 287

Qy        342 DYIKEALMKIEDVYSVEDLNIWSLTSGKSTAIVHIQLIPGSSSKWEEVQSKANHLLLNTF 401
                ::: |: :| | ::  |:||:||  :   |||:   : :  : |   |:  |   |
Db        288 TAVRDLLLSVEGVEALHSLHIWALTVAQPVLSVHIAIAQNTDA--QAVLKTASSRLQGKF 345

Qy        402 GMYRCTIQLQSYRQEVDRTCANCQSSS 428
              :  |||:: | ::: : |  ||  |
Db        346 HFHTVTIQIEDYSEDM-KDCQACQGPS 371
```

```
Search completed: July 29, 2004, 17:43:09
Job time : 57 secs
```

Blank sheet